ABSTRACT

Title of Document: DEMAND-RESPONSIVE AIRSPACE SECTORIZATION AND AIR TRAFFIC CONTROLLER STAFFING

Shin-Lai Tien, Doctor of Philosophy, 2010

Directed By: Professor Paul M. Schonfeld
Department of Civil and Environmental Engineering

This dissertation optimizes the problem of designing sector boundaries and assigning air traffic controllers to sectors while considering demand variation over time. For long-term planning purposes, an optimization problem of clean-sheet sectorization is defined to generate a set of sector boundaries that accommodates traffic variation across the planning horizon while minimizing staffing. The resulting boundaries should best accommodate traffic over space and time and be the most efficient in terms of controller shifts. Two integer program formulations are proposed to address the defined problem, and their equivalency is proven. The performance of both formulations is examined with randomly generated numerical examples. Then, a real-world application confirms that the proposed model can save 10%–16% controller-hours, depending on the degree of demand variation over time, in comparison with the sectorization model with a strategy that does not take demand variation into account.

Due to the size of realistic sectorization problems, a heuristic based on mathematical programming is developed for a large-scale neighborhood search and implemented in a parallel computing framework in order to obtain quality solutions within time limits. The impact of neighborhood definition and initial solution on heuristic performance has been examined. Numerical results show that the heuristic and the proposed neighborhood selection schemes can find significant improvements beyond the best solutions that are found exclusively from the Mixed Integer Program solver's global search.

For operational purposes, under given sector boundaries, an optimization model is proposed to create an operational plan for dynamically combining or splitting sectors and determining controller staffing. In particular, the relation between traffic condition and the staffing decisions is no longer treated as a deterministic, step-wise function but a probabilistic, nonlinear one. Ordinal regression analysis is applied to estimate a set of sector-specific models for predicting sector staffing decisions. The statistical results are then incorporated into the proposed sector combination model. With realistic traffic and staffing data, the proposed model demonstrates the potential saving in controller staffing achievable by optimizing the combination schemes, depending on how freely sectors can combine and split. To address concerns about workload increases resulting from frequent changes of sector combinations, the proposed model is then expanded to a time-dependent one by including a minimum duration of a sector combination scheme. Numerical examples suggest there is a strong tradeoff between combination stability and controller staffing.

DEMAND-RESPONSIVE AIRSPACE SECTORIZATION AND AIR TRAFFIC CONTROLLER STAFFING

By

Shin-Lai Tien

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2010

Advisory Committee:

Professor Paul M. Schonfeld, Chair

Professor Michael O. Ball

Associate Professor David J. Lovell

Assistant Professor Lei Zhang

Dr. Robert L. Hoffman

Acknowledgements

In the journey of pursuing my doctorate I was blessed to have the opportunities to work with many great researchers. My sincere gratitude goes to my advisor Professor Paul Schonfeld for his guidance and encouragement throughout my study at UMD. His enthusiasm in research motivated me to keep the passion and faith in discovering research problems and exploring methods.

I would like to thank Professor Michael Ball for providing me support academically and financially. He always devoted his time to me when I needed help and provided insightful comments on my research problems. I greatly appreciate every discussion with him and am deeply affected by his research attitude and philosophy.

I would like to thank Professor David Lovell for providing me guidance on research projects. He encouraged me to think 'outside of the box' and to approach a problem from different angles.

I owed a lot to Dr. Robert Hoffman, who has spent countless hours on discussing my random ideas, helping me identify possible research directions. His selfless personality and positive attitude to life and work makes him a role model to me. Also, thanks to him and people at Metron Aviation for sponsoring data collection.

I would like to thank Professor Lei Zhang for serving in my dissertation committee and sharing his research ideas to help me seek possible application areas of my work.

Thanks to my friends in the NEXTOR office, Andy Churchill, Kleoniki Vlachou, Charles Glover, Alex Nguyen, Moein Ganji, Nasim Vakili, and Prem Swaroop, who

made my stay enjoyable. Every lunch outing was full of joy and life experience sharing, and every conference we have attended was memorable.

My oversea study was not alone because of the company of my friends from Civil Engineering Department, Ning Yang, Shiaaulir Wang, Frank Chen, Jason Chou, Lisa Chen, Ming Chen, Xiaodong Zhang, Zichun Li, Pei-Wei Lin and Yue Liu.

I would also like to offer my regards and blessings to all of those who helped me in any respect during the completion of this dissertation: Arash Yousefi from Metron Aviation, Stephanie Chung from FAA, Edward Huang from U. C. Berkely, Michael Drew, Michael Bleom, and Shanon Zelinski from NASA Ames, and David Gianazza from ENAC.

Most of all, I would like to express my deepest appreciation to my family. My wife and parents made everything possible in my life. Their sacrifice and unconditional support have driven me forward. They gave me love and hope so that I would be able to achieve this life goal and expect many great things to come.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Research Background and Motivation

By 2025, traffic in the U.S. airspace is expected to double or triple. In response to expected growth in air travel demand, the Next Generation Air Transportation System (NEXTGEN) is designed to revolutionize vehicle performance, navigational technology and air traffic management (ATM) concepts. Amongst the advanced ATM concepts being developed, dynamic airspace configuration (DAC) addresses the need to dynamically allocate both air traffic control (ATC) resources and the airspace structure to meet real-time demand profiles (Kopardekar et al., 2007). The overarching assumption is that if the ATC resources and airspace are more efficiently utilized and designed (i.e. to reduce controller requirements or workload), demand fluctuations can be accommodated, which means more traffic will be served, thereby reducing air traffic control (ATC) delays to airspace users.

Current enroute airspace over continental U.S. is divided into 20 control areas, each of which is managed by an Air Route Traffic Control Center (ARTCC), or "center" for short. The control area of each center is further divided into sectors, which are the basic subdivisions for enroute air traffic control. Fig. 1-1 illustrates how the airspace is divided into centers and then sectors.

Air traffic controllers who serve sector traffic provide safe separation between aircraft by giving maneuvering instructions to pilots. On request, they also provide other information such as weather and turbulence conditions. For staffing and managerial needs, each center may have 6 to 8 areas of operation, each of which might consist of

6 to 9 sectors. Air traffic controllers have to familiarize with and obtain certification for the sectors in the area they will be serving.

The design of current sectors is based on historical traffic patterns, jet routes, location-specific restrictions, and controller workload considerations. It has evolved over a long time based on incremental addition of new technologies and procedures for air traffic control. Sector boundaries might stay fixed for months if there is no significant change of traffic patterns.

For day-to-day operation, resectorization due to demand fluctuation regularly occurs when sectors are combined or split based on traffic demand. Sectors with less traffic are combined with others to save controller resources whereas sector with heavy traffic may be split in order for controllers to deal with temporary overloads. Because expertise in pattern recognition does not develop quickly, enroute controllers take an average of about 3 years to be certified as Certified Professional Controller (CPC). (Stein et al., 2006) Resectorization upon a given airspace structure preserves some stability in the controllers' pattern cognition. From a human factors perspective, controllers' strategy options for routing and conflict resolution will not be severely limited by such resectorization activities.

Various mathematical methods of enroute airspace partitioning or sectorization have been proposed in the literature (as reviewed in Chapter 2) for designing airspace sectors with appropriate controller workloads and addressing selected design objectives, e.g. alignment with traffic flow or buffering for aircraft maneuvering. A question of particular interest for sector boundary design that has not been answered

is how temporal variation of demand impacts the controller workability of airspace and the efficiency of staffing.

Surprisingly, efficient controller staffing does not receive much attention although unlimited controller staffing would be an expensive policy. In the U.S., controller labor costs have increased from $82.98 per flight in FY1998 to $137.81 per flight in FY2006 (FAA, 2006). The Federal Aviation Administration plans to hire and train more than 15,000 controllers over the next decade, in response to controller attrition and an anticipated increase in air travel (FAA, 2007). There is a strong need in the U.S. to take controller costs into account when designing sector boundaries or management plans.

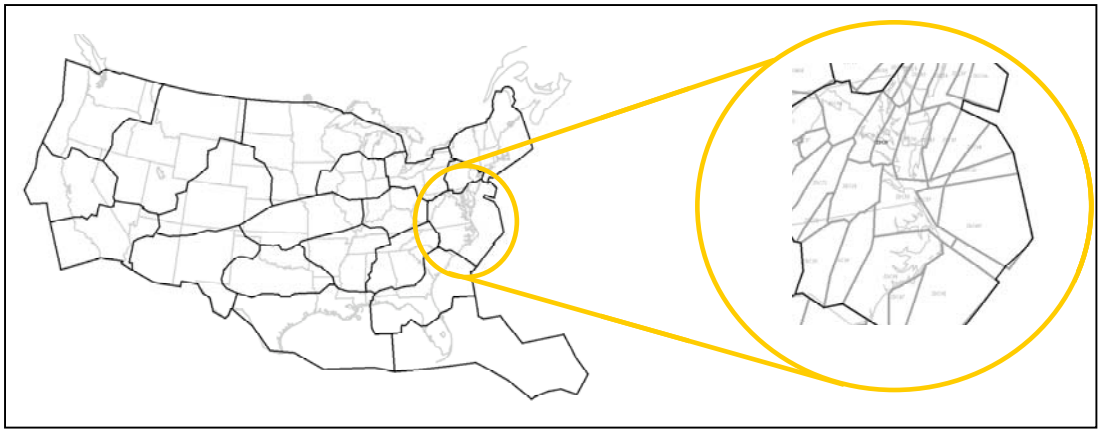


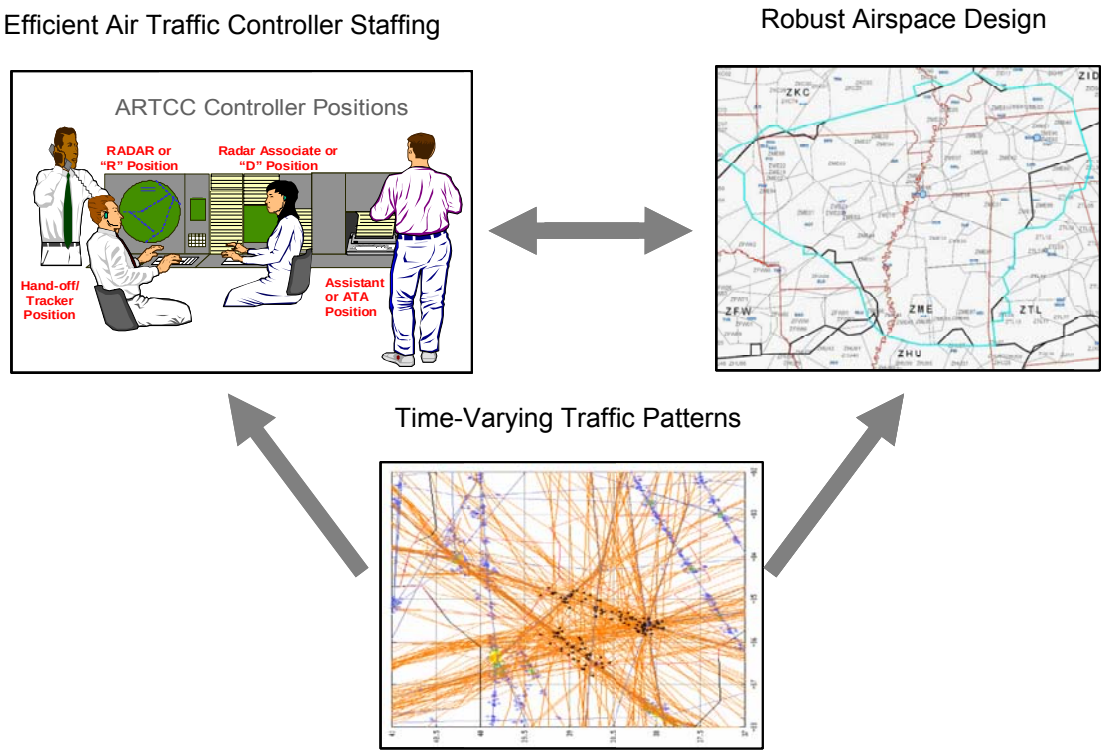
Figure 1-1 Structural Airspace Control Area

1.2 Relation among Airspace Design, Traffic and Controller Staffing

Airspace design, traffic patterns, and enroute airspace resources, i.e. human controllers, interact. It is expected that airspace design has to best accommodate daily traffic and balance controller workload among sectors. However, under current practice, sector boundaries do not change for months or years. Thus, traffic variation over time and space should be properly addressed in airspace design models.

A generally adopted philosophy in airspace design is to evaluate and balance controller workload among sectors. Controller workload is an expression of how air traffic control activities are perceived by human controllers, and it heavily depends on their experience levels and individual differences. To be measurable objectively, it can be further decomposed into three main categories: monitoring, conflict resolution, and coordination (Delahaye, 1994). Various traffic complexity metrics in these three categories are thus proposed to objectively assess controllers' workload level. For example, perhaps the most intuitive metric is the number of aircraft handled in a sector per given time interval, which is directly associated with the controllers' monitoring workload. In addition, the control area and shape of a sector are also related to the controller workload in aircraft handoff and coordination with neighboring sectors. Poor sector geometry would not only increase the frequency of aircraft handoffs but also impact the maneuver ability to resolve conflicts. It is safe to say that the combination of sector boundaries and aircraft trajectories jointly determines and is closely related to controller workload.

Particularly in the U.S., when sector traffic is high, a common way to deal with increasing workload is to assign multiple controllers to serve a sector. An example of sector capacity estimation based on controller staffing is given in Fig. 1-3, although the controller staffing issue has not yet been linked to the airspace design problem.



Number of Controllers by Function and Number of Aircraft Worked

Function	Number of Aircraft Worked During 15-Minute Interval	Number of Controllers
High Altitude Radar Sector	0	0
	1 - 12	1
	13 - 17	2
	18 - 29	3
	30+	4
Low/Both Altitude Radar Sector	0	0
	1 - 10	1
	11 - 14	2
	15 - 24	3
	25+	4
A-Side	0	0
	1 - 73	1
	74+	2

NOTE: For application, count aircraft worked for radar sector controller positions during current 15-minute interval. Count aircraft worked for A-side positions at +30 minutes from current 15-minute interval.

Figure 1-3 Effects of Number of Controllers on Sector Capacities (FAA, 1997)

1.3 Research Scope and Objectives

This dissertation addresses the treatment of demand variation in airspace sector design and management problems. This dissertation is intended to answer several near-term or mid-term questions in DAC study of NextGen and assumes the availability of certain new navigational and automation technology, e.g. data link, automated conflict detection with proposed resolutions, ADS-B[1], etc. Optimization models and solution methodologies are proposed for incorporating newly introduced design concepts into airspace configuration and sector management, namely time-varying demand patterns and efficient controller staffing. It is proposed to include within airspace configuration models the interaction among traffic, airspace design, and controller staffing. Specifically, when designing sector boundaries, traffic variability is treated by varying the number of controllers working each sector. This capitalizes on the existing practice of multi-controller teams. The outputs of the proposed sectorization model provide optimized sector boundaries as well as a least-cost staffing assignment (i.e. number of controllers assigned to each sector throughout the day). Previous studies either ignore this capacity control mechanism or are in conflict with it by assuming a uniform number of controllers per sector. The subject of dynamic airspace configuration will be approached from two perspectives.

Strategically, a model of clean-sheet sectorization is intended to generate a set of boundaries that accommodates traffic variation across the planning horizon. The

[1] Automatic Dependent Surveillance-Broadcast (ADS-B) is a system that uses precise location data from the global satellite network and enables both pilots and controllers to see radar-like displays with highly accurate, real time traffic data from satellites. The system will also give pilots access to weather services, terrain maps and flight information services. The improved situational awareness will mean that pilots will be able to fly at safe distances from one another with less assistance from air traffic controllers. (FAA, 2008)

resulting sector boundaries will best accommodate traffic over space and time and be the most efficient in terms of controller shifts. The focus is on demand variations that occur throughout the day, rather than over the course of months or years.

Tactically, under given sector boundaries, a common way to deal with traffic demand variation over time and space is to temporarily combine sectors with low traffic or to assign more controllers to work busy sectors. An optimization model is proposed to create an operational plan for dynamically combining or splitting sectors and assigning controllers in order to support 24/7 operation. Such a sector combination/split plan should minimize controller shifts and satisfy certain operational rules, such as minimizing changes of control regions over successive periods. Since this tactical model focuses on real-time response, controller workload will increase nonlinearly with the number of aircraft worked. Thus, statistical analysis is applied to quantify the relation between controller staffing and selected traffic metrics, and its results are then incorporated into the sector combination problem for predicting sector staffing decisions.

1.4 Dissertation Organization

The organization of this dissertation is described below and the relation among chapters is illustrated in Fig. 1-4.

- Chapter 1 provides the motivation and defines the research scope and objectives.
- Chapter 2 reviews the state-of-the-art in modeling airspace design and quantifying controller workload. The design concepts and methods as well as deficiencies will be compared and discussed.

- Chapter 3 formally defines the problem of sectorizing clean-sheet airspace while considering multi-period demand patterns and time-varying controller staffing. Two integer programming formulations are proposed, and their equivalency is proved. A branch-and-cut algorithm is proposed to optimally solve one of the formulations. Numerical examples with realistic traffic data are analyzed to demonstrate the soundness of the model and examine its coherence with current practice.
- Chapter 4 develops a heuristic based on mathematical programming for finding quality solutions within time limit. The heuristic is implemented in a parallel computing framework so that multiple neighborhoods are examined simultaneously. Several schemes for selecting promising neighborhoods have been tested and their results are compared. The sensitivity of the initial solutions is also analyzed.
- Chapter 5 quantifies the relation between staffing decisions and traffic. Historical staffing and traffic data are processed, and their statistical relation is then estimated. The effect of observation period duration is also examined. The estimation results are a crucial component for the sector combination model which links sector staffing directly to sector capacity.
- Chapter 6 proposes a model for supporting daily center operations that provides sector combination/split decisions and minimizes controller shift usage. The model determines how a sector should be combined or split from neighboring sectors and what staffing level is required to serve the traffic.

- Chapter 7 summarizes the main contributions of this dissertation and suggests extensions for future work.

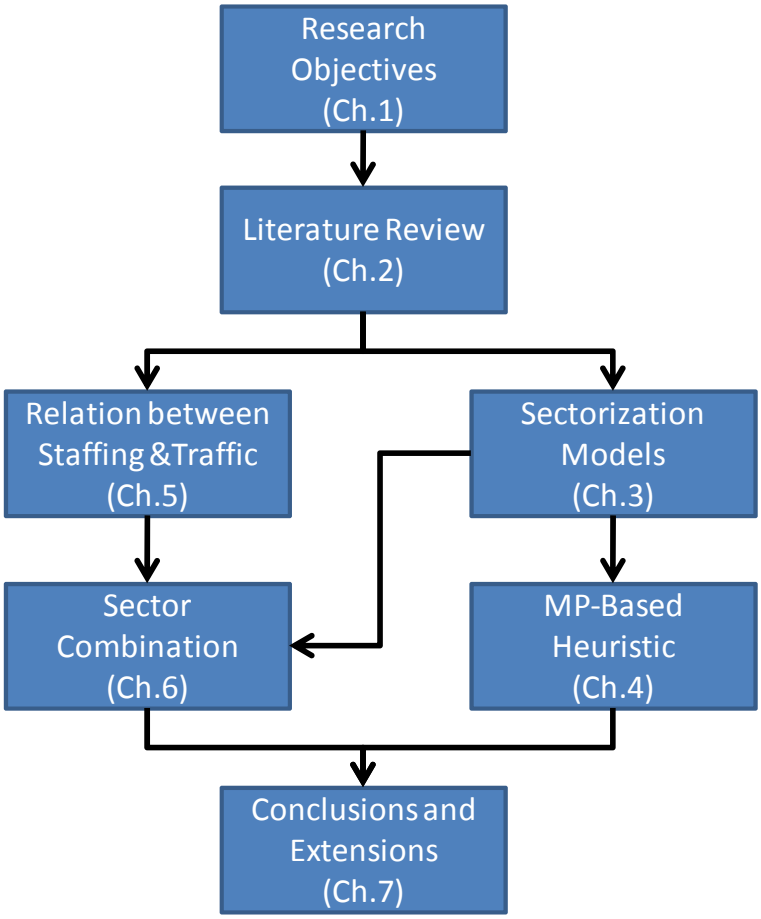


Figure 1-4 Dissertation Chapter Interrelation

Chapter 2: Prior Research on Sectorization and Sector Combination Problems

This chapter provides literature review on three subjects that are closely related to the objectives of this dissertation: 1) clean-sheet airspace partitioning; 2) construction of the relations between controller staffing and traffic complexity metrics; 3) design of sector combination/split schemes.

2.1 Methods for Clean-Sheet Sectorization

Clean-sheet sectorization is one particular subject in airspace design, which focuses on designing radar-controlled sectors for unstructured airspace, i.e. ignoring current sector boundaries and air routes. It serves as a cornerstone for dynamic airspace configuration (DAC) research since one of the DAC research areas is dynamic adjustment of control areas in response to real-time traffic. Various methods and models have been explored to address numerous design criteria, e.g. sector design should afford optimum flight profile procedures that enable flights to reach desired altitudes, optimum speeds, and climb/descent rates without interruption for ATC operational or organizational reasons. Human controller workability on designed boundaries is also an important consideration as most of existing studies take into account controllers' perception of traffic complexity and try to balance workload among sectors. Other (secondary) sector design considerations might include equipment and spectrum constraints (e.g. radio coverage), local boundary constraints (e.g. reserved airspace), aircraft performance mix, shallow-angle boundary crossings,

and room for controller maneuvering of aircraft (e.g. keeping intersection points away from the boundary).

In the FAA's Order of Facility Operation and Administration, a set of factors is listed for consideration in determining the size and configuration of enroute sectors, such as traffic volume and flow, location and activity of terminals, special operations/procedures, coordination requirements, radar/radio coverage, equipment limitations, and airway alignments. (FAA, 2010) Workload should also be distributed equitably among sectors. More desirably, sector boundaries should contain the longest possible segments of airways and align with the primary traffic flow, so the coordination workload between sectors can be reduced.

However, these sector design criteria often interact or conflict with one another, so it would be ambitious to meet all objectives in one modeling effort. In order to implement the designed airspace in practice, Conker et al. (2007) proposed a framework in Fig. 2-1 for clean-sheet airspace design, consisting of three important aspects: 1) airspace partitioning, 2) controller workability evaluation, and 3) sector boundary evaluation and improvement. Since the practicability of designed airspace depends on various factors, their research pointed out the need to have each aspect addressed by an individual module separately developed so that favorable modeling techniques can be applied in order to incorporate more realistic or comprehensive design concerns.

While researchers address various geometric considerations in their Airspace Partitioning models, workload balancing is one universal concern. Delahaye et al. (1994a, 1998) proposed genetic algorithms to obtain sectors with well-balanced

workload. Other design considerations include minimum sector transit time, sector shape convexity. Trandac and Duong (2003) also considered workload balancing and proposed a two-phase approach using graph partitioning heuristics to find an initial sectorization plan and employing constraint programming techniques to locally optimize based on a set of geometric constraints.

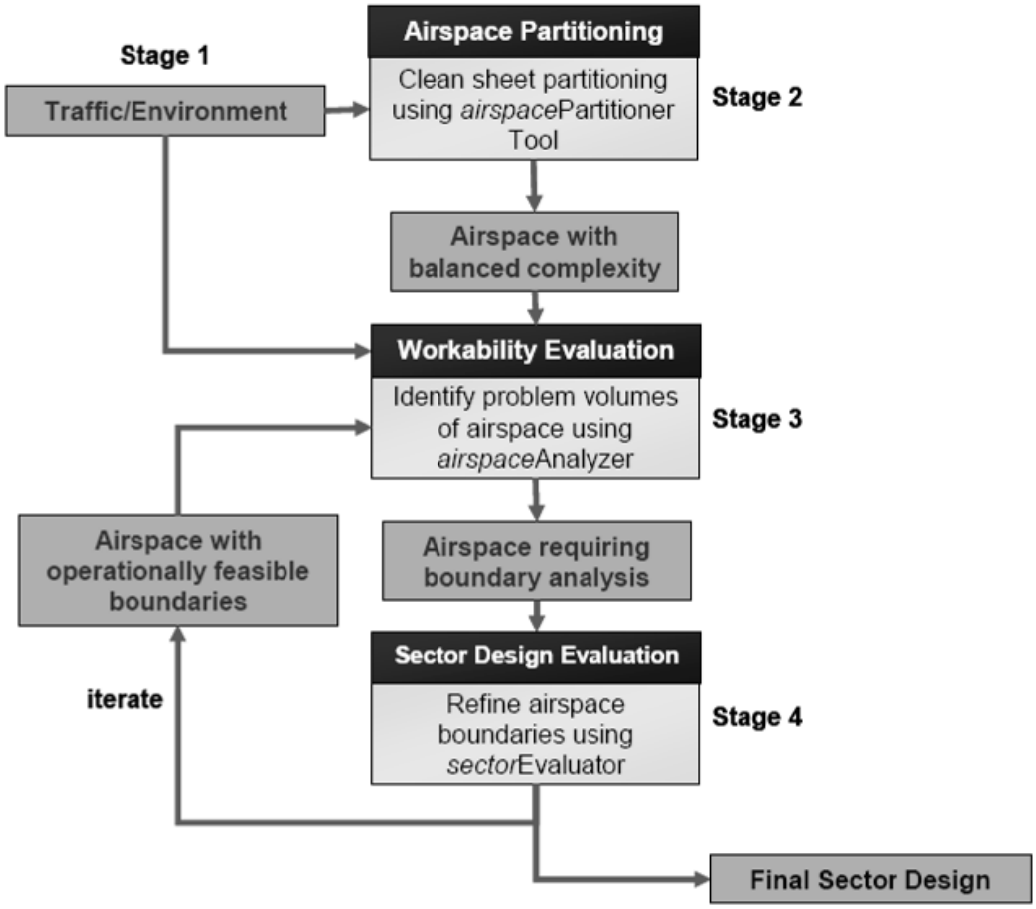


Figure 2-1 Conker et al.'s Clean-Sheet Airspace Design and Evaluation Process

For sector design purposes, Klein (2005) proposed an algorithm for determining a center boundary by iteratively combining hex-cells in order to achieve well-balanced workload among resulting centers. The number of radar hits in a given area serves as a traffic mass metric, and its high correlation to controller workload is demonstrated

in a simulation environment. The proposed method, which resembles a bacteria population growth model, gradually expands the size of centers in the selected locations. To be applicable to sector boundary design, the model has to address the relation between traffic flow characteristics and sector shape.

A prominent methodology in sectorization is the application of computational geometry (CG) algorithms. The primary advantage of CG over classical optimization techniques is that virtually any computable design criteria can be incorporated. In Basu et al. (2008), airspace is partitioned into sectors of well-balanced workload with the workload metric approximated by aircraft count. Mitchell et al. (2008) incorporated additional metrics on workload variation into CG algorithms to capture traffic variation: peak, average, transfer flight counts. Sector shape concerns are addressed by limiting the aspect ratio of resulting sectors. Sabhnani et al. (2010) then extends the bisection algorithm in Basu et al. (2008) by considering the abstraction of major traffic flows into generating geometric cuts that partition the airspace. Xue (2008) applied optimization algorithms to improve the solutions from a Voronoi Diagram[2] solution in order to meet various objectives in airspace design.

Another optimization approach formulates a sectorization problem as an integer program over a discretized airspace. Unlike the CG approach, this tile-and-group idea decomposes the target airspace into small hexagons and then groups them through an optimization model. A critical consideration is to insure that all the small hexagons grouped into a sector are connected.

[2] Given *n* generating points in a plane, the problem of a Voronoi Diagram is to partition the plane into *n* convex polygons such that each polygon contains exactly one generating point and every point in a given polygon is closer to its generating point than to any other.

In Yousefi (2004), the hexagonal cells used to tile airspace are clustered into a pre-determined number of sectors, constrained by a workload balancing criterion. Each hexagon is associated with a calculated workload metric and considered as a customer with positive demand in an Uncapacitated Facility Location (UFL) problem. An MIP model based upon UFL assigns customers (hexagons) to the facilities (sector) of a pre-determined number. The primary concern in the objective function is to balance the workload amongst sectors. However, connectedness is not guaranteed for any feasible solution to his formulation.

A later mixed integer program (MIP) formulation proposed by Yousefi et al. (2007) assures connectedness by building upon a network flow model. The target airspace is tiled with equal-sized hex-cells. A network is constructed by representing hexagons as nodes and their adjacency as links. The amount of workload is then seen as commodity in this network flow problem. Flow conservation constraints maintain that the resulting sectors are connected components. The link cost of underlying network is defined as the crossings of flights between two adjacent hexagons. By minimizing the link cost, the model addresses the concern of aligning the sector shape with the major flight traffic. The workload balancing consideration among sectors is imposed as constraints. Later, Drew (2008) proposed to set identical costs on both directions of a link in order to avoid bizarre sector shapes. This modification creates visually promising sector shapes, and a smoothing method is then applied to remove the zigzag boundaries due to hexagon combinations.

2.2 Opportunities for Improving Airspace Design Models

Despite all the techniques that have been tried, sectorization techniques in the literature heavily emphasize the balancing of workloads across sectors – a shared objective in almost all the models. The idea is that prevention of sector overloads will reduce the need for controllers to apply enroute flow restrictions. But workload balancing tacitly presumes that sectors have (or should have) equivalent capacity across the planning horizon. To the contrary, sector capacity varies with the number of controllers working that piece of airspace, as illustrated in Chapter 1.2. An enroute sector in the United States is managed by a team of up to four controllers. Therefore, the capacity of a given sector may be treated as a variable, increasing in steps with the number of controllers assigned to that sector.

Another common feature found in previous studies is the design of airspace for forecasted workload aggregated over one planning horizon (e.g. one day or week). However, this is too coarse in time to capture the traffic variations that occur throughout the day. Some of the techniques can accommodate peak workload, but then the sectorization tends to cater to those peaks rather than to the variance. If demand variation is not considered, the resulting boundary design might end up using controller resources inefficiently.

A conceivable convenient approach might be to reapply a static sectorization method as frequently as needed. For instance, if traffic demand in the 12:00 – 14:00 time period were significantly different than the demand pattern in the 10:00 – 12:00 time period, then a resectorization would be proposed at 12:00. However, this practice is

highly disruptive to controller workflow. Frequent and drastic changes in boundary definition will increase the intensity of workload and the chance of operational errors. In practice, it might take at least several minutes for controllers to "get the picture". It is especially difficult for controllers to make staffing or boundary changes during intense activity. Jung, et al (2010) conducted a human-in-the-loop experiment on the effect of sector boundary changes on air traffic controllers and observed that there is a 12.7% increase in average workload due to frequent boundary changes. Moreover, a controller has to be familiar with (and certified on) sector boundaries in his or her area of operation (FAA, 2010). All these concerns favor some degree of stability in sector boundaries.

Although temporary adjustment of sector control areas (e.g. combining quiet sectors or splitting busy sectors) is commonly seen in practice, the potential sector boundaries, once determined, will last for months or years. In practice, significant sector boundary changes are made only in the hours when traffic levels are low. For the foreseeable future, as long as human controllers still play the major role in traffic control and coordination, it is reasonable to assume that, wholesale resectorization during "the heat of battle" will remain impractical. When designing sector boundaries, there is a strong need to account for traffic patterns over time.

2.3 Relations among Controller Staffing, Controller Workload and Traffic Complexity Metrics

An enroute controller has all three basic functions for a sector: ATC-to-pilot communication, data processing and management (e.g. handling flight strips), and coordination with other air traffic controllers. When traffic is slow, one controller can

perform these tasks. However, during high demand periods, these responsibilities are often spread among multiple controllers. A common configuration is to augment the primary Radar (R-side) controller working a given sector with a second controller (Data or D-side). Fig. 2-2 shows the activities of traffic vs. controller staffing throughout a typical busy day at the ZNY center. The yellow line depicts the traffic handled in each quarter hour. Each stacked bar from bottom to top shows the number of sectors in ZNY that are currently assigned 0, 1, 2, and 3 controller positions. During the daily peak, more sectors tend to have multiple controllers.

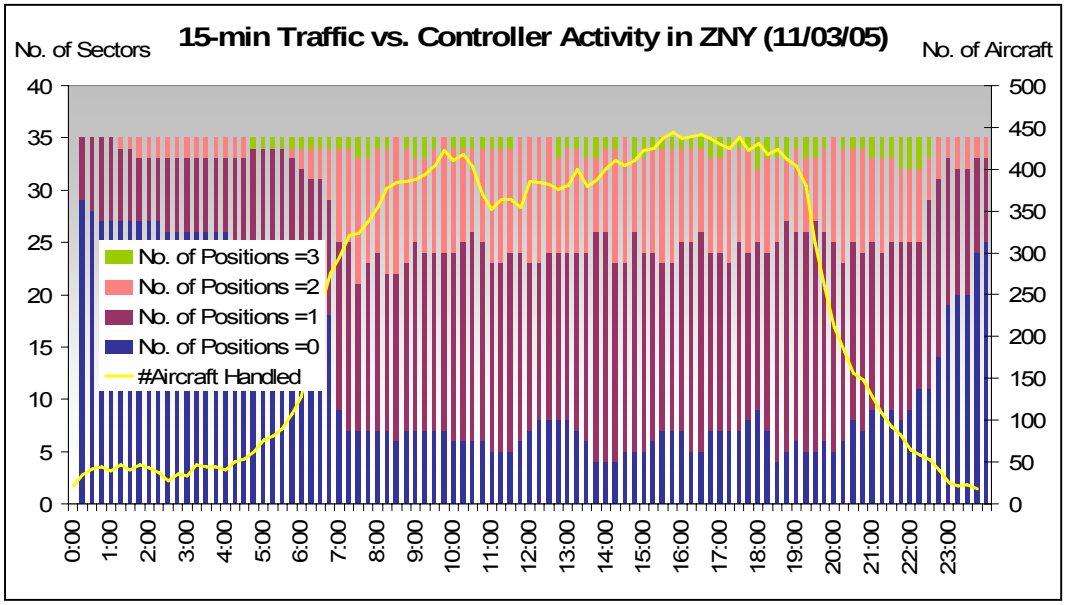


Figure 2-2 An Illustration of Sector Controller Staffing throughout a Day

For the purpose of this dissertation, the relation between controller staffing and traffic is essential in the sense that sector staffing is directly related to sector capacity, thus determining the upper bound of workload in a sector. To approximate such staffing decisions directly from traffic characteristics, Tien and Schonfeld (2008) used a probit regression model to estimate the probability of multiple controller positions given aircraft count and sector characteristics. Their model was used to estimate

annual controller requirements at a national level, so their selection of covariates included a limited range of complexity metrics, that is, sector aircraft count, traffic density and sector area.

A probit regression model is suitable for predicting categorical, ordinal response variables, such as using $1,\ldots,N$ controllers per sector. Denoting x as the measurable factors and ε as the unobservable error factor, it is grounded in a latent regression specified as:

$$y^* = \mathbf{x}\boldsymbol{\beta} + \varepsilon \ ,$$

where y* is unobserved and ranges from $-\infty$ to ∞, $\boldsymbol{\beta}$ is the coefficient vector of the covariate $\mathbf{x}$, and ε is the error term.

Based on certain $\mathbf{x}$, we want to know which response category will be chosen. Since the latent variable y^* is unobservable, what we do observe from the data is the staffing decision n. Thus, the following relation is assumed:

$$n = \begin{cases} 1 & \text{if } -\infty \le y^* \le \mu_1 \\ 2 & \text{if } \mu_1 \le y^* \le \mu_2 \\ \ldots & \\ N & \text{if } \mu_{N-1} \le y^* \le \infty \end{cases} ,$$

where $\mu_1,\ldots,\mu_{N-1}$ are the threshold values between ordinal categories.

If the error term ε belongs to the normal distribution, the probabilities of decision categories can be estimated using the covariate values in the measurement equation and by taking the inverse of the normal distribution function. For instance, Prob($n \le 1$) = $\Phi(\mu_1 - \mathbf{x}\boldsymbol{\beta})$. The probabilities for individual categories P_n can then be derived by

taking the differences of the cumulative probabilities for the groups in order, e.g.

$$P_1 = \Phi(\mu_1 - \mathbf{x\beta}),\ P_2 = \Phi(\mu_2 - \mathbf{x\beta}) - \Phi(\mu_1 - \mathbf{x\beta}),\ \ldots,\ P_N = 1 - \Phi(\mu_{N-1} - \mathbf{x\beta}).$$

In other words, the probability for the first category is the first cumulative probability; the probability for the second category is the second cumulative probability minus the first; and so on. Then the prediction can be made by choosing the category with highest probability. The implication of the probability structure is shown in Figure 2-3. Each categorical probability is actually the area under the CDF within the range defined by threshold values (Long, 1997).

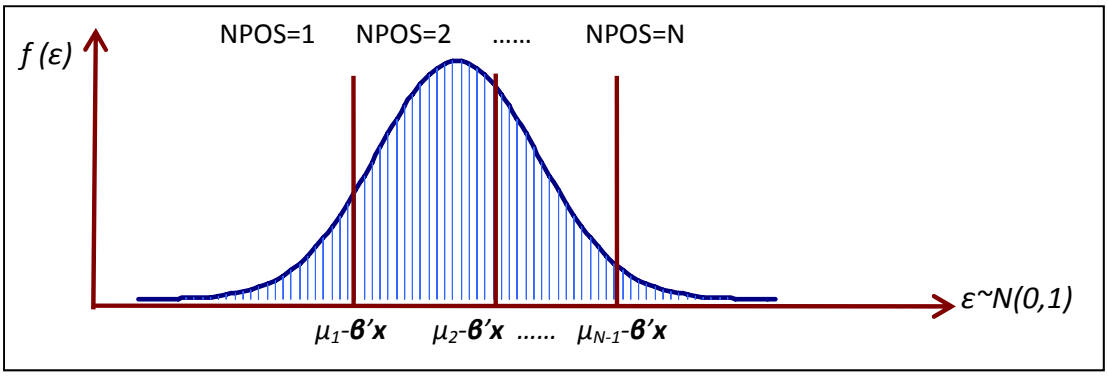


Figure 2-3 Illustration of Probit Model Structure

The selection of the covariates for the probit regression model should address several aspects of controller workload. Controller workload is a combined result of many contributing factors of traffic and airspace characteristics, which is illustrated in Fig. 2-4 from Mogford et al. (1995). From Section 2.1, it can be seen that studies in airspace design always require translating traffic characteristics into a few quantifiable metrics so that those metrics can be utilized by researchers to model various design concepts, determining desirable boundaries of control areas.

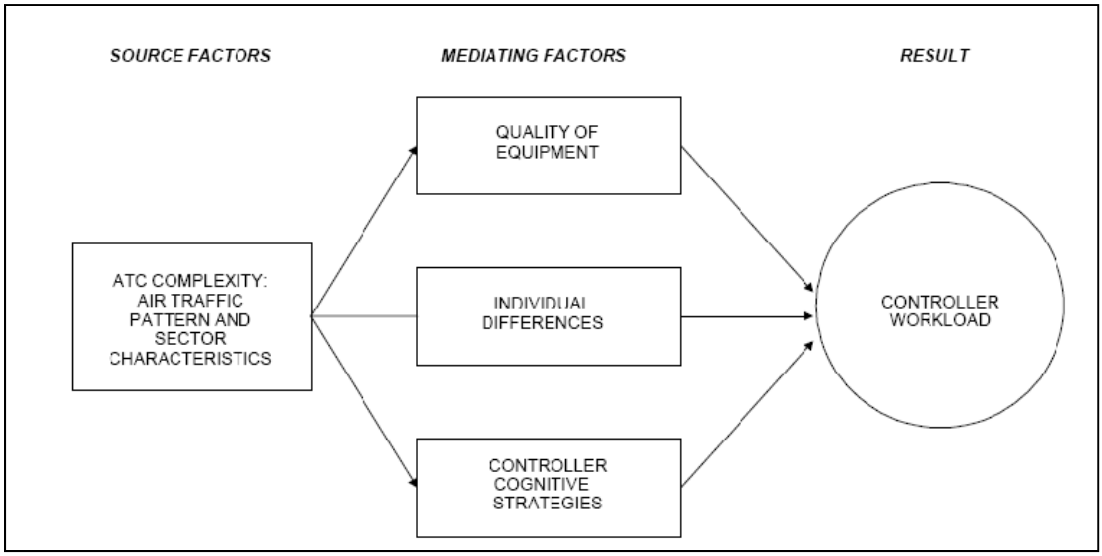


Figure 2-4 Factors Contributing to Controller Workload by Mogford et al. (1995)

To objectively quantify the relation between workload and traffic complexity, researchers have defined various metrics, which can be roughly categorized into three aspects (Mogford, 1994):

- Physical aspects of the sector/airspace structure, e.g. sector volume, number of flight levels, number of intersecting flight paths, etc.
- Air traffic movement, e.g. number of cruising, ascending/descending, or transitioning aircraft.
- Combination of the above two, e.g. sector aircraft density and average sector transit time.

From the perspective of controllers' tasks, Delahaye et al. (1994) decomposed the controller workload into three categories:

- Conflict resolving workload relates to all the actions to solve traffic conflicts.
- Coordination workload relates to information exchange among controllers of adjacent sectors and the pilots of aircraft crossing sector boundaries.

- Trajectory monitoring workload relates to the workload of continuously checking trajectories of aircraft in a sector.

These identified categories will help researchers choose suitable traffic complexity metrics to better describe the instantaneous operational environment and to approximate controller staffing decisions. A comprehensive list of metrics from Mogford, et al. (1995) is listed In Table 2-1. The selection of metrics depends on the statistical significance of estimated models.

Table 2-1 List of Candidate Complexity Metrics (Mogford, et al., 1995)

1. Number of aircraft
2. Aircraft density or traffic volume
3. Aircraft handled in prior time interval (e.g. last hour)
4. Number of arrivals
5. Number of departures
6. Number of emergencies
7. Number of special flights
8. Coordination
9. Traffic mix (arrivals, departures, and overflights)
10. Number of airport terminals
11. Traffic distribution
12. Staffing
13. Weather conditions
14. Equipment status
15. Number of communications with aircraft
16. Number of communications with other sectors
17. Presence of conflicts
18. Number of path changes
19. Preventing conflicts (crossing or overtake)
20. Number of handoffs and printouts
21. Handling pilot requests
22. Traffic flow structure
23. Clustering of aircraft
24. Control adjustments involved in merging and spacing
25. Mixture of aircraft types
26. Climbing and descending aircraft
27. Number of intersecting flight paths
28. Number of required procedures
29. Number of military flights
30. Airline hub location
31. Weather and its severity
32. Aircraft routing

33. Special use airspace
34. Sector geometry
35. Sector size
36. Requirements for longitudinal and lateral spacing
37. Radar coverage
38. Frequency congestion
39. Number of altitudes used

2.4 Sector Combination Problem

Sector combination problem is defined as the combination/split of sectors in response to traffic variation. For enroute airspace, consolidation of sectors adapts resources to changes in demand. Sector combination is commonly used during periods of light traffic such as night-time flow shifts. When two sectors are combined into one, a determination is made that the two sectors need to be combined to balance the workload of air traffic controllers. The choice is procedural and initiated by an area supervisor. When sector traffic is high and creates datablock clutter on the radar display, sectors may be split vertically to manage both the aircraft count per sector and ease the datablock clutter (Lee, et al., 2008).

Bloem et al. (2009) summarize the feedbacks of subject matter experts and mentioned that under current air traffic control environment, the stability and controllers' familiarity of the sector combinations would be the main concerns to the area supervisors. On the other hand, there are the benefits expected from sector combination/splitting activities. From managerial viewpoints, combining sectors would reduce the staff required to manage a piece of airspace and lead to fewer airspace-induced flight restrictions, e.g. more direct routings. When splitting sectors, workload is generally reduced and thus safety is increased. With less sector traffic,

controllers could provide higher quality services to aircraft, such as weather information, direct routings, and altitude changes to reduce turbulence.

Sector combinations responsive to traffic do promote the efficiency of ATC resources (e.g. controllers), although dynamically changing sector combinations over time raises concerns of practicability. It is expected that with new automation technology introduced in NextGen (data link, automated conflict detection with proposed resolutions, ADS-B, etc.) the need for familiarity with sector combination could be alleviated, and extra burden to controller workloads from sector combination would not be the same as it is today (Gupta et al., 2009).

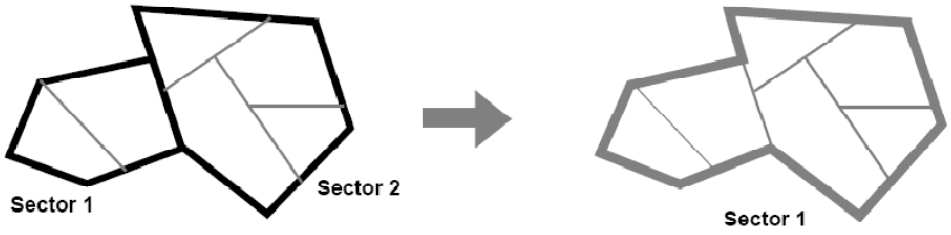
In practice, a busy sector might need multiple controllers to provide service, or possibly some of its fixed posting areas (FPA) might be designated to adjacent sectors in order to deal with controller overload. An FPA is a three-dimensional volume of airspace and can be considered as a fundamental unit of airspace. The airspace of a sector is a set of one or more contiguous FPAs that constitute a specified sector. An FPA has a default sector but may be designated to others due to ATC operational needs. Meyers et al. (1998) described possible types of combination based on sectors or FPAs.

Sector-to-sector combination assigns inactive sector to one of its active neighbors. It is also possible that more than two sectors are combined. Fig. 2-5(a) depicts before and after the combination of Sectors 1 and 2. (The bold, gray line is sector boundary after combination. The thin, gray line is the boundary of FPA.)

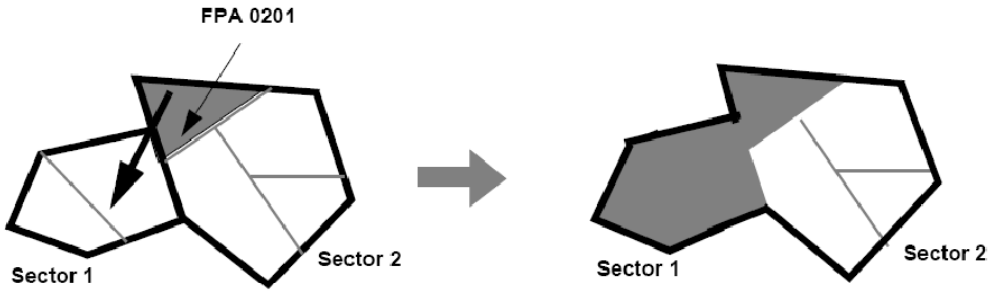
FPA-to-sector combination designates an FPA to another sector. It is commonly used for dealing with sector overload. In Fig. 2-5(b), a FPA of Sector 2 is combined to

Sector 1. The shaded area after combination is then the new airspace controlled by controllers serving Sector 1.

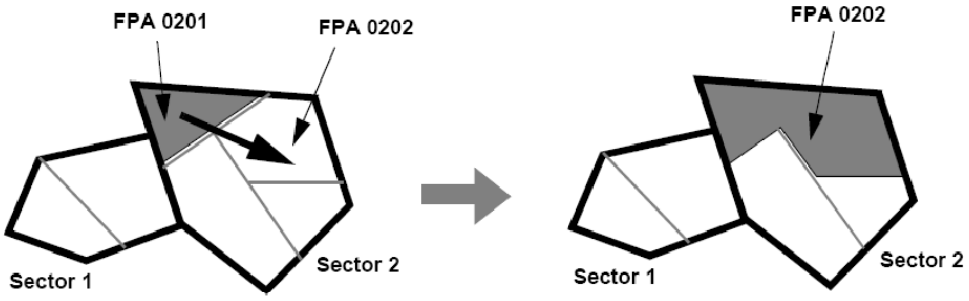
FPA-to-FPA combination combines two adjacent FPAs. In Fig. 2-5 (c), two FPAs of Sector 2 are combined. However, this type of assignment of one (or more) FPA to another FPA could occur in two cases: 1) all FPAs are in the same sector, as indicated above, or 2) one FPA is in one sector and another FPA is in another sector.



(a) Sector-to-Sector Combination



(b) FPA-to-Sector Combination



(c) FPA-to-FPA Combination

Figure 2-5 Typical Types of Sector Combination (Meyers et al., 1998)

The sector combination problem is studied under a common assumption that traffic patterns are given. Its goal is to minimize the usage or promote the utilization of air traffic control resources. However, the Air Traffic Flow Management (ATFM) flow restrictions are the ultimate actions to maintain the safety and minimize delay, which makes the sector combination problem a less important concern. Thus, the assumption of no adjustment to traffic patterns seems to be reasonable for now for the purpose of the sector combination problem.

The sector combination problem appeared as early as in Delahaye et al. (1995). They solved an abstract version of the problem with genetic algorithm. The design objectives were balancing workload after combination and minimizing coordination workload. The modeling effort in their work illustrated the importance of maintaining connectivity and handling combined sector workload. Sector capacity was assumed to have a maximum value for combined sectors.

Verlhac and Machon (2001) considered the minimum duration of a sector configuration (i.e. the layout of resulting combined sectors). They proposed integer programming models that searched for a suitable set of pre-defined layouts to best accommodate time-varying demand by minimizing sector capacity deficits. Since a layout has a minimum duration, sector demand was allowed to temporarily overload. Their model assumed one position for one sector and was meant for planning two days before the day of operations.

Gianazza, et al. (2002a) and Gianazza and Alliot (2002b) proposed a cost function for optimizing sector combination by weighting two design objectives, i.e. the number of

resulting sector control positions and the deficit and surplus of sector capacity. The resulting configuration yields a set of sector groups, each of which have demand as close as possible to capacity. Due to the problem complexity, three solution search techniques were proposed and their performances were compared.

Later, by training an artificial neural network model with historical data, Gianazza (2007) was the first in the literature to use realistic measures of traffic complexity to predict the probability of a sector being merged, manned, or split. Then, possible airspace configurations, no longer limited to a set of pre-defined ones, were enumerated and evaluated to find the best configuration that met design objectives similar to those proposed in Gianazza et al. (2002a). Their predictions were claimed to be realistic from the traffic manager's perspectives.

Gianazza (2008, 2009) incorporated smoothed traffic complexity metrics in order to address the issues of traffic variability and to improve model predictability. By comparing the number of sectors in the predicted configuration with that in the actual one, it was found that the proposed model performed better with the metrics smoothed over 30 minutes.

More recently, the sector combination problem has received attention from the researchers in the U.S. Bloem et al. (2008, 2009) assumed sector Monitor Alert Parameter (MAP) values as complexity and capacity surrogates[3] and proposed a heuristic algorithm that combined adjacent sectors by shortening sector capacity gaps after combination. The algorithm was tested with simulated traffic data to demonstrate its potential in increasing sector utilization rate. Drew (2009) then

[3] MAP is a traffic complexity metric that jointly considers flight counts and average flight dwell time in a sector. (FAA, 2010)

developed an optimization version of Bloem et al's problem by formulating a variant of the multi-commodity network flow problem.

Although the approaches applied in the literature are different, common objectives are to balance workload among resulting sectors or to minimize the sector count. None of the existing studies links this problem to efficient controller staffing, which implies that staffing levels required for daily operations are left to the discretion of area supervisors.

2.5 Summary

This chapter summarizes prior studies on airspace sectorization techniques and ways of combining sectors to achieve operational efficiency. Controller staffing has not been addressed in any existing models, which yields a research need for incorporating this factor into the design of sector boundaries and sector combination schemes. In addition, existing models deal with single-period demand patterns, which motivates this study to address traffic variation over time. In the following chapters, these research directions will be pursued, and corresponding optimization models and solution techniques will be proposed.

Chapter 3: Integer Programs for Optimal Sectorization

This chapter is organized as follows. Section 3.1 states the problem of airspace sectorization while considering multi-period demand patterns and time-varying controller staffing. Sections 3.2 and 3.3 propose two integer programming formulations, and discuss their theoretical properties. Section 3.4 proposes a solution algorithm that dynamically identifies violated constraints for one of the formulations. Section 3.5 contains the numerical experiments on randomly generated data for analyzing the solution quality as well as comparing computational performance of both formulations. The soundness of the models and their coherence with current practice will be demonstrated and examined. In Section 3.6, a real world example is solved with the proposed sectorization objective and one commonly used in the literature, respectively. The advantages and weakness of both objectives will be discussed. Conclusions and main contributions will be summarized in Section 3.7.

3.1 Problem Statement and Complexity

The approach taken here for airspace sectorization is first to tile the airspace with equal-sized hexagonal cells (applicable to other polygonal shapes) and then to determine how to group those cells to achieve design objective(s). With this tile-and-group approach, a graph can be constructed by setting each hex-cell as a node and the adjacency of two nodes as a link, as illustrated in Fig. 3-1. The main assumptions made here are:

- Traffic patterns, which can be historical, typical or forecasted, are known a priori.

- Sector demand and sector capacity levels are measured in the same units, e.g. number of aircraft, radar hits, or a composite metric estimated from traffic complexity analysis during a certain period.
- Traffic demand is additive across cells. (This will be elaborated in the computation experiment section.)

As mentioned in Chapter 2, sector design concerns are too many to be included in one single model, so in this complexity analysis we limit the focus to a generic problem – grouping cells so that multi-period demands are served and controller resources are minimized. Here is a general description of the underlying problem:

Sectorization with Multi-period Demand Patterns and Variable Sector Capacity Choices (MPVC):

INSTANCE: Graph $G = (V, E)$, weights $d(v,t) \geq 0$ for each node $v \in V$ at each time $t \in T$. A set of possible choices of capacity values k_r, where $r \in R$. A step-wise increasing functional relation that maps those capacities to their associated costs c_r. A positive integer Q. For $i = 1,...,m$, denote $K(V_i,t)$ and $C(V_i,t)$ the capacity choices and associated costs for the subset of nodes V_i at time t, where $K:(V_i,t) \to \{k_r \mid r \in R\}$ and $C:(V_i,t) \to \{c_r \mid r \in R\}$.

QUESTION: Is there a partition of V into at most m disjoint, connected subgraphs $V_1,...,V_m$ such that $\sum_{v \in V_i} d(v,t) \leq K(V_i,t)$ for $1 \leq i \leq m$ and $t \in T$, and such that the total cost is no greater than Q, i.e. $\sum_{t \in T} \sum_{i=1}^{m} C(V_i,t) \leq Q$?

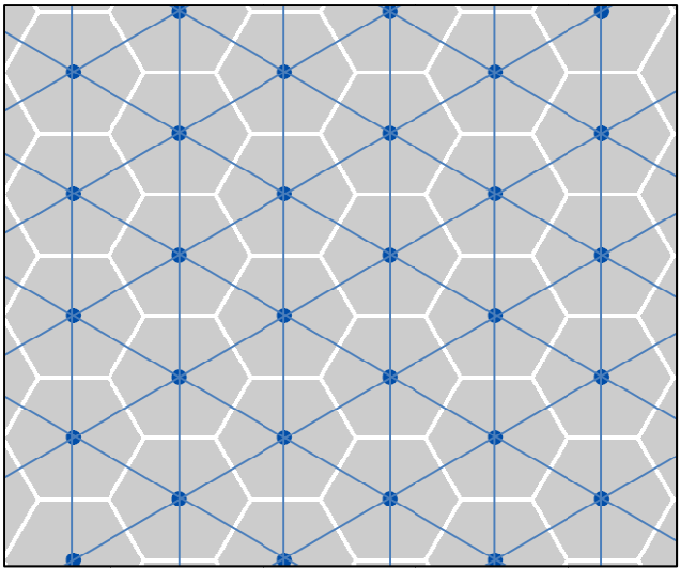


Figure 3-1 Illustration of Tile-and-Group Approach

Theorem 3.1: MPVC is NP-Complete.

Proof:

We prove this by a reduction from the Partition Problem in Garey and Johnson (1979):

Partition Problem (PP): Given a finite set A and a size $s(a) \in Z^+$ for each $a \in A$. Is there a subset $A' \subseteq A$ such that $\sum_{a \in A'} s(a) = \sum_{a \in A \setminus A'} s(a)$?

We transform PP into MPVC. Given an instance of IPP, a graph can be constructed in polynomial time: Consider each element in A as a node in a graph and two dummy nodes v_s and v_e. The node set V and edge set E of a graph can be set as follows and shown in Fig. 3-2:

- $V = \{v_s, 1, 2, ..., |A|, v_e\}$
- $E = \{(v_s, 1), (v_s, 2), ..., (v_s, |A|), (1, v_e), (2, v_e), ..., (|A|, v_e)\}$

Set $T = \{1\}$, $R = \{1\}$, and let $C(V_i,t) = 1$, $K(V_i,t) = \sum_{a \in A} s(a)/2$, $\forall i,t$, and $Q = 2$. For all $t \in T$, set node weights $d(v_s,t) = d(v_e,t) = 0$ and $d(a,t) = s(a)$ for $a = 1,...,|A|$, a reduced instance of MPVC from PP has thus been constructed.

It is to be shown that there is a "yes" solution to an instance of Partition Problem (IPP) if and only if there is a "yes" solution to the reduced instance of MPVC (IMPVC).

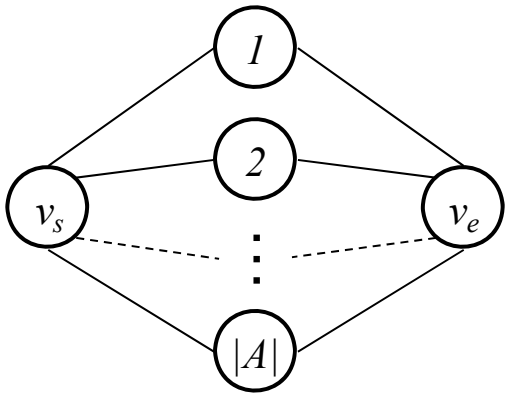


Figure 3-2 Fig. 3-2 Graph Constructed from Instance of Partition Problem

→ If there is a 'yes' solution to IPP , we need to show that there is a partition that satisfies the capacity constraints with total cost less or equal than Q. From the solution of IPP, the corresponding nodes in A' along with node v_s constitutes a node set, say V_1, with total node weight of $\sum_{a \in A'} s(a)$; the other nodes in $A \setminus A'$ along with node v_e constitutes a node set V_2 with total node weight of $\sum_{a \in A \setminus A'} s(a)$. The solution of IPP ensures $\sum_{a \in A'} s(a) = \sum_{a \in A \setminus A'} s(a) = \sum_{a \in A} s(a)/2$. Thus, for $i = 1,2$, the total node weights of V_i is less or equal to the capacity provided, i.e.

$\sum_{v \in V_i} d(v,t) = \sum_{a \in A} s(a)/2 \le K(V_i,t)$, and total cost is $\sum_{t \in \{1\}} \sum_{i \in \{1,2\}} C(V_i,t) = 2 \le Q$. This is a 'yes' solution to IMPVC.

← If there is a 'yes' solution to IMPVC , there are exactly two subsets of nodes V_1 and V_2 , each of which has the sum of node weights $\sum_{v \in V_i} d(v,t)$ for $i = 1,2$, and $V = V_1 \cup V_2$. Suppose V_1 contains v_s. Set $A_1 = V_1 \setminus \{v_s\}$ and $A_2 = V_2 \setminus \{v_e\}$ for IPP. Since $V_1 \cup V_2 = V$, $A_1 \cup A_2 = A$, which implies that either A_1 or A_2 contains half or more of the total weight. Without loss of generality, say A_1. That is, $\sum_{a \in A_1} s(a) \ge \sum_{a \in A} s(a)/2$.

In addition, $\sum_{a \in A_1} s(a) = \sum_{v \in V_1} d(v,t) \le K(V_1,t) = \sum_{a \in A} s(a)/2$.

Thus, $\sum_{a \in A_1} s(a) = \sum_{a \in A} s(a)/2$. Since $\sum_{a \in A_1} s(a) + \sum_{a \in A_2} s(a) = \sum_{a \in A} s(a)$, it is concluded that $\sum_{a \in A_2} s(a) = \sum_{a \in A_1} s(a)$. This is a 'yes' solution of IPP.

In the next two sections, two integer programming formulations will be proposed to find optimal sector boundaries across the planning horizon while considering efficient controller staffing for individual periods and to include the following features:

1. Equal sized hex-cells are grouped into several connected subgraphs with minimum controller costs. This is related to the feasibility of the cell grouping techniques. In Yousefi (2005), the proposed formulation imposes additional constraints to increase the tendency of creating connected sectors, e.g. a cell and one of its neighbors have to be in the same sector, but does not entirely eliminate the possibilities of sectors disconnected from the solution space.

Yousefi et al. (2007) reformulate the problem based on network flow formulation and thus maintain the connectivity of any resulting sector.

2. The size (capacity) of each connected subgraph (sector) and associated cost increases in steps with controllers assigned. This feature accounts for the step-wise nature of controller staffing.
3. The shape off sectors (the contour of hex-cells in the connected subgraphs) preferably aligns with major traffic flows in order to maintain a reasonable flight dwell time, reducing the frequency of aircraft handoff and its associated workload.

3.2 The Spanning Tree-Based Formulation P_{STB}

A spanning tree is a 1-edge-connected graph, which means if an edge is removed all the nodes no longer form a connected component. The spanning tree formulation P_{STB} presented here guarantees the connectivity of each resulting sector by maintaining a spanning tree, whose size (capacity) will be determined through optimization.

Preliminaries are first given: Each hex-cell is represented as node $i \in \{1,...,I\}$. Let $\delta(i)$ be the set of nodes adjacent to i, and the edge set $E = \{(i,j) \mid i = 1,...,I, j \in \delta(i), i > j\}$ describes hex-cell adjacency relation. Denote the index $p \in \{1,...,P\}$ as choices of controller positions (i.e. sector capacity values), $t \in \{1,...,T\}$ as time period, and $k \in \{1,...,K\}$ as generic sector. Here a sector k has no physical meaning and serves as a potential sink to receive the demands assigned to it. Whether a sector k is used will be determined by the optimization model.

Decision variables used in this formulation are defined as follows:

$$x_i^k = \begin{cases} 1, & \text{if node } i \text{ is assigned to sector } k. \\ 0, & \text{otherwise.} \end{cases}$$

$$y_{ij}^k = \begin{cases} 1, & \text{if edge}(i, j) \text{ is assigned to sector } k. \\ 0, & \text{otherwise.} \end{cases}$$

$$z_p^{k,t} = \begin{cases} 1, & \text{if position } p \text{ of sector } k \text{ is staffed at time } t. \\ 0, & \text{otherwise.} \end{cases}$$

In Theorem 3.1, we describe a generic version for airspace sectorization with multi-period demand patterns, whose only objective is to minimize the controller cost over the planning horizon. While airspace design usually involves more than one design objective, here we start with two cost terms that should be minimized in order to address this multiple-objective concern.

- Objective 1 – determine efficient controller usage by minimizing total controller hours $\sum_{k,p,t} h_p z_p^{k,t}$, where h_p is the cost of controller position p.
- Objective 2 – align sector shape with major traffic flows by minimizing the total edge cost $\sum_{(i,j)\in \mathrm{E},k} c_{ij} y_{ij}^k$, where c_{ij} is the cost of edge (i, j).

In the first objective, controller resources can be capitalized by setting the cost h_p related to controller head count or other monetized values to reflect wages, experience levels, etc.

Another sector design criterion addressed here is that sector shapes should align with the major traffic flows. Controllers have indicated that this is a paramount concern in sector design. Flow alignment reduces workload in the form of handoffs. To the extent possible, sector boundaries should minimize coordination needs and promote

overall system flexibility to support user-preferred trajectories. To achieve this goal, the second objective is to minimize the cost of all edges that are assigned to sectors. The edge cost c_{ij} is used as a surrogate to describe the handoff traffic between two adjacent hex-cells i and j. It can be set as a monotonically decreasing function of traffic crossings in between, so it creates a tendency to combine hex-cells with heavy coordination needs.

A binary integer program is formulated as follows:

$\mathrm{P}_{\mathrm{STB}}$:

$$\min f(x,y,z) = \mu \sum_{k,p,t} h_p z_p^{k,t} + \sum_{(i,j)\in \mathrm{E},k} c_{ij} y_{ij}^k$$

Subject to:

$$\sum_k x_i^k = 1 \quad \text{for all } i \in \{1,...,I\} \tag{3-1}$$

$$y_{ij}^k \le \begin{cases} x_i^k \\ x_j^k \end{cases} \quad \text{for all } (i,j) \in E, k \in \{1,...,K\} \tag{3-2}$$

$$z_{p+1}^{k,t} \le z_p^{k,t} \quad \text{for all } k \in \{1,...,K\}, t \in \{1,...,T\}, p \in \{1,...,P-1\} \tag{3-3}$$

$$\sum_i d_i^t x_i^k \le \sum_p U_p z_p^{k,t} \quad \text{for all } k \in \{1,...,K\}, t \in \{1,...T\} \tag{3-4}$$

$$\sum_{(i,j)\in E} y_{ij}^k = \sum_{i\in\{1,...,I\}} x_i^k - z_{p=1}^{k,t=1} \quad \text{for all } k \in \{1,...,K\} \tag{3-5}$$

$$\sum_{(i,j)\in E(X)} y_{ij}^k \le |X| - 1 \quad \text{for any } X \subset \{1,...,I\}, |X| \ge 2, k \in \{1,...,K\} \tag{3-6}$$

$$x, y, z \in \{0,1\}$$

For the objective function, if there is a strong tradeoff between controller staffing cost and the edge cost, i.e. flight alignment cost, the number of resulting sectors exceeds what is actually needed. As a result, more controller positions are employed and the capacity offered is not well utilized. Since the number of sectors is not pre-determined and controller cost is the primary objective to be optimized, we want the tradeoff between two objectives as minimal as possible. The multiplier μ will determine the dominance of controller cost for this multi-objective problem. In the numerical experiment section we will discuss how a proper value of μ should be selected.

Constraint (3-1) is the node assignment constraint, which requires each node to be assigned to exactly one of the sectors. Constraint (3-2) describes the node-edge relation, preventing an edge from being assigned to a sector if either of its end nodes is assigned to that sector. Constraint (3-3) describes the step-wise increasing nature of capacity resulting from adding controller positions. It also ensures that additional positions are not staffed before earlier ones, e.g. the second position is not used until the first one is used, and so on.

Let d_i^t denote the demand (or weight) associated with node i at time t and assume a non-trivial case in which $d_i^t > 0$. Denoting U_p the capacity added by using position p, Constraint (3-4) ensures that at each time period, the sum of node weights in a sector cannot exceed the capacity determined by controller staffing. Note that $z_{p=1}^{k,t} = 0$ means that the first position of sector k is not staffed at t, implying that $\sum_i d_i^t x_i^k = 0$, i.e. $x_i^k = 0$ for all $i \in \{1,...,I\}$ and thus that sector k is not "used" for all

the t's. Thus, without loss of generality, we let $z_{p=1}^{k,t=1}$ serve as an indicator of the usage of sector k.

Constraints (3-5) and (3-6) are adopted from the polytope representation of a spanning tree problem, which has been proven by Edmonds (1970) to be as follows:

Theorem: Given a connected undirected graph G, $n := |V(G)|$, the spanning tree polytope is:

$$\left\{ x \in [0,1]^{E(G)} : \sum_{e \in E(G)} x_e = n-1, \sum_{e \in E(G[X])} x_e \le |X| - 1 \; for \; \varnothing \ne X \subset V(G) \right\},$$

whose vertices are integral and exactly the incidence vectors of spanning trees of G.

In Constraint (3-5), since we do not know a priori how many nodes a "used" sector would have, the variable of the first controller position $z_{p=1}^{k,t=1}$ serves as an indicator of whether sector k is "used" for all the time periods. If $z_{p=1}^{k,t=1} = 1$, then the number of assigned edges is equal to the number of assigned nodes less one; otherwise, both sides of the equation are thus zero. Constraint (3-6) is called a ***Cycle Elimination Constraint***, which says that for any subset of nodes, the number of edges assigned to a sector cannot exceed the cardinality less one. The number of constraints of this type is exponential, so it would be impractical to enumerate all the possibilities a priori. Constraints (3-5) and (3-6) jointly construct a tree structure for the edges assigned to a sector so that all the nodes that are assigned to the same sector will be connected.

3.3 The Network Flow-Based Formulation P_{NFB}

The main characteristics of the formulation P_{STB} are that it directly assigns nodes and edges to sectors and it uses spanning tree properties to ensure the connectivity of nodes assigned to the same sector. With the same objectives as in P_{STB}, an equivalent formulation P_{NFB} is constructed as a variant of the network flow problem with side constraints.

The underlying graph of the formulation P_{NFB} is similar to P_{STB}: Each cell corresponds to a node in the network and is denoted as $i \in \{1,...,I\}$. Let $\delta(i)$ be the set of nodes adjacent to i. There will be two links directed from/to each of its neighbor nodes, and thus the edge set E in P_{NFB} is defined as: $E = \{(i, j) \mid i = 1,...,I, j \in \delta(i)\}$. Also denote the index $p \in \{1,...,P\}$ as choices of controller positions (i.e. sector capacity values), and $t \in \{1,...,T\}$ as a time period.

The demand at node i in period t is denoted $d_i^t \geq 0$, which is the workload (e.g. aircraft count) to be served at hex-cell i. Conceptually, each unit of demand can be considered as a flow commodity on the network flow problem. The commodity originating from a node will flow along the network through a series of nodes to exactly one sink node, i.e. the commodity here is non-bifurcated. The sink node of a commodity is, however, unknown a priori. The choice of sink nodes will be determined through the optimization process. In this setting, each sink node from the optimization results represents the existence of a sector. All nodes contributing to that sink belong to that sector. These node-to-sink (or cell-to-sector) relations are

determined by tracing paths in post-optimization processing. The boundaries of a sector are then depicted by the contour of hex-cells grouped in that sector.

In order to represent the choices of sector capacity values, a special treatment made here is that we augment the graph by attaching a dummy link to each node. In Fig. 3-3, during period t, the flow merged at node i either goes to one of its neighboring nodes or passes to its dummy link. The maximal amount of commodities a dummy link can carry at t is step-wise increasing and will be set as a function of additional controllers.

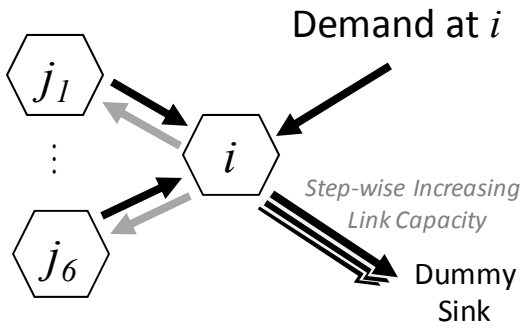


Figure 3-3 Node Augmentation for P_{NFB}

The continuous decision variable $u_{ij}^t \geq 0$ represents at period t the amount of commodities carried by link $(i, j) \in E$, and $s_i^t \geq 0$ represents at period t the amount of commodities (sinking at node i) carried by the dummy link of node i. The binary decision variables of P_{NFB} are the link variable w_{ij} and the controller variables $g_{i,p}^t$, which are defined as follows:

$$w_{ij} = \begin{cases} 1, & \text{if link } (i, j) \text{ carries any positive flow commodities.} \\ 0, & \text{otherwise.} \end{cases}$$

$$g_{i,p}^t = \begin{cases} 1, & \text{if position } p \text{ of node } i \text{ is staffed at period } t. \\ 0, & \text{otherwise.} \end{cases}$$

In particular, the controller variable $g_{i,p}^t$ in $\mathrm{P_{NFB}}$ represents the capacity values of the dummy link through the following functional relation:

$$\text{The capacity of dummy link at node } i = \sum_{p=1}^{P} U_p g_{i,p}^t$$

where U_p is the capacity increased by adding position p.

The mixed integer program is formulated as follows:

$\mathrm{P_{NFB}}$:

$$\min f(u,s,w,g) = \mu \sum_{i,p,t} h_p g_{ip}^t + \sum_{(i,j)\in E} c_{ij} w_{ij}$$

Subject to:

$$\sum_{j\in\delta(i)} u_{ji}^t + d_i^t = \sum_{j\in\delta(i)} u_{ij}^t + s_i^t \quad \text{for all } i \in \{1,...,I\}, t \in \{1,...,T\} \tag{3-7}$$

$$g_{i,p+1}^t \le g_{i,p}^t \quad \text{for all } i \in \{1,...I\},\ p \in \{1,...,P-1\},\ t \in \{1,...,T\} \tag{3-8}$$

$$s_i^t \le \sum_{p=1}^{P} U_p g_{i,p}^t \quad \text{for all } i \in \{1,...I\},\ t \in \{1,...,T\} \tag{3-9}$$

$$\sum_{j\in A_i} w_{ij} + g_{i,p=1}^{t=1} = 1 \quad \text{for all } i \in \{1,...I\} \tag{3-10}$$

$$u_{ij}^t \le \left(\sum_p U_p \right) w_{ij} \quad \text{for all } i \in \{1,...,I\}, j \in \delta(i), t \in \{1,...,T\} \tag{3-11}$$

$$u, s \ge 0, g, w \in \{0,1\}$$

Constraint (3-7) is a flow conservation constraint. At node i and time t, the commodity (or demand) originated at i and received from some of the adjacent nodes j's should sink at node i or pass to its neighbors.

Constraint (3-8) ensures the step-wise increasing nature of controller staffing, e.g. the second controller (data position) may be needed when the workload of the first controller (radar position) exceeds a certain threshold.

When a node is determined as a sink node, the commodity merged at it will be further directed to its dummy link. Constraint (3-9) requires that the amount of commodity that sinks in node i at time t cannot exceed the capacity offered at its dummy link, defined by the staffing decision at time t. Thus, if there are commodities sinking in node i at time t, then we know at least $g_{i,p=1}^{t}=1$ because of the step-wise increasing relation described by Constraint (3-8).

Constraint (3-10) ensures that the commodity sinking at node i is not split, i.e. it is either conducted to one of the neighboring nodes or to the dummy link. Given that w_{ij} is directional and has no time index, if none of the links adjacent to node i is chosen, i.e. $\sum_{j\in A_i} w_{ij}=0$, then the dummy link at time t will have to carry s_i^t commodities. Without loss of generality and assuming that for such node i, s_i^t is positive for all $t\in\{1,...,T\}$, the first position must be used for all the time periods, which implies $g_{i,p=1}^{t=1}=...=g_{i,p=1}^{t=T}=1$ and node i is a sink node. Thus, $g_{i,p=1}^{t=1}$ can be used as an indicator of whether node i is chosen as a sink node.

Constraint (3-11), which is a typical Big-M constraint in the network design problem, sets the upper bound of the flow of commodities carried on link (i,j). The tightest upper bound is $\sum_{p=1}^{P} U_p$, which is the maximum amount workload a sector can accommodate. If $w_{ij}=0$, then $u_{ij}^{t}=0$ for any t.

The objective function of P_{NFB} is equivalent to that of P_{STB}. The first term is the controller cost, and the second is the link cost (or flow alignment penalty), where $c_{ij} = c_{ji}$. It is worth noting that in a typical network design problem, e.g. Hochbaum and Segev (1989), the link cost can be further categorized into the fixed and variable link costs. While the fixed link cost considers a fixed value for a chosen link, the variable link cost weights the cost of a chosen link by how much flow it carries. To capture the flow pattern, Yousefi et al. (2007) chose to minimize in the objective function the variable link cost, instead of the fixed link cost. There is no obvious advantage of choosing either fixed or variable cost in the formulation since both choices are expected to favor the combination of two nodes with strong crossing traffic.

We first prove a property of P_{NFB} solution in Lemma 3-1 and then the equivalency of P_{NFB} and P_{STB} in Lemma 3-2.

Lemma 3-1: Given symmetric link cost, i.e. $c_{ij} = c_{ji}$, at the optimality of P_{NFB} the links that forms a sector is a minimum link-cost spanning tree of the nodes in that sector.

Proof:

Recall that if a node r has $g_{r,p=1}^{t=1} = 1$, it is called a sink node. By Constraint (3-10), if a node is not determined to be a sink node by the optimization model, only one outbound link will be selected to carry positive flow. Thus, a sink node and all the links that carry flow to it form a connected subnetwork. The property of unsplittable flows makes this subgraph a spanning tree, and the sink node can be seen as the root

of the tree.

Assume (u^*, s^*, w^*, g^*) is an optimal solution to P_{NFB} and $f^*_{NFB} = \sum_{i,p,t} h_p g^{*t}_{ip} + \sum_{(i,j)\in E} c_{ij} w^*_{ij}$ is the optimal objective function value. For any tree T_r with a root node r, if there exists an link (i, j) where both end nodes i, j are in the tree T_r and the link variable $w^*_{ij} = 0$ and if the link cost $c_{ij} \le c_e$ where $w^*_e = 1$ and e is an edge on the i-j-path in T_r, then setting $w_{ij} = 1$ and $w_e = 0$ would yield an objective function value $(f^*_{NFB} - c_e + c_{ij}) \le f^*_{NFB}$. Thus, (u^*, s^*, w^*, g^*) cannot be an optimal solution, which contradicts the assumption.

Therefore, at the optimality of P_{NFB}, the links that form a sector constitute a minimum link-cost spanning tree of the nodes in that sector.

Lemma 3-2: Given symmetric link cost, i.e. $c_{ij} = c_{ji}$ in P_{NFB}, a solution (u, s, w, g) is feasible (optimal) to P_{NFB} if and only if there is a solution (x, y, z), which is feasible (optimal) to P_{STB}.

Proof:

($P_{NFB} \rightarrow P_{STB}$)

Given a solution of (u, s, w, g) to P_{NFB}, we can know the number of sectors K' from the cardinality of the sink nodes i, such that $s^t_i > 0$ for any $t \in \{1, ..., T\}$. Label those i's with $l_1, l_2, \ldots l_{K'}$. From the P_{NFB} solution, we can identify the trees $T_1, T_2, ..., T_{K'}$ associated with $l_1, l_2, \ldots l_{K'}$.

A corresponding solution (x, y, z) to $\mathrm{P}_{\mathrm{STB}}$ can thus be constructed from a solution (u, s, w, g) to $\mathrm{P}_{\mathrm{NFB}}$ as follows:

For each sector index $k \in \{1, ..., K'\}$,

- Set node assignment variable $x_j^k = 1$ if j is in tree T_k; otherwise, $x_j^k = 0$.
- Set edge assignment variable $y_{ij}^k = 1$ if the link variable $w_{ij} = 1$ and i, j are both in the tree T_k; otherwise, $y_{ij}^k = 0$.
- Set controller variable $z_p^{k,t} = 1$ if $g_{l_k, p}^t = 1$; otherwise, $z_p^{k,t} = 0$.

Constraints (3-3) and (3-4) are automatically satisfied because they are equivalent to Constraints (3-8) and (3-9) in $\mathrm{P}_{\mathrm{NFB}}$. The property proven in Lemma 3-1 ensures the solution of $\mathrm{P}_{\mathrm{NFB}}$ is a tree, so the constraints related to tree construction, i.e. Constraints (3-1), (3-2), (3-5) and (3-6) are satisfied.

($\mathrm{P}_{\mathrm{STB}} \rightarrow \mathrm{P}_{\mathrm{NFB}}$)

Recall that the controller variable $z_{p=1}^{k,t=1}$ serves as the indicator of whether sector k is used. Given a solution of (x, y, z) to $\mathrm{P}_{\mathrm{STB}}$, the number of used sectors is $K' = \sum_{k=1}^{K} z_{p=1}^{k,t=1}$.

To distinguish used and unused sectors, assume $z_{p=1}^{k,t=1} \le z_{p=1}^{k+1,t=1}$, so $z_{p=1}^{k,t=1} = 1$ for $k = 1, ..., K'$, and $z_{p=1}^{k,t=1} = 0$ for $k = K'+1, ..., K$.

From the $\mathrm{P}_{\mathrm{STB}}$ solution, denote as T_k the set of tree nodes associated with sector $k \in \{1, ..., K'\}$, where $T_k = \left\{ i \mid x_i^k = 1, \text{ for all } i = 1, ..., I \right\}$. A corresponding solution

(u, s, w, g) to $\mathrm{P}_{\mathrm{NFB}}$ can thus be constructed as follows:

For each $k \in \{1,...,K'\}$,

- Designate an arbitrary node $\tilde{i}$ in T_k as the root node.
- Set $g^t_{\tilde{i},p} = 1$ if $z^{k,t}_p = 1$; otherwise, $g^t_{i,p} = 0$ for $i \in T_k$, $i \neq \tilde{i}$.
- From the $\mathrm{P}_{\mathrm{STB}}$ solution, there is a unique path from each node in T_k to the root node $\tilde{i}$. Denote as $succ(i)$ the immediate successor of $i \in T_k$ on the path to $\tilde{i}$. Set $w_{ij} = 1$ if $j = succ(i)$; otherwise, $w_{ij} = 0$.
- For each $w_{i,succ(i)} = 1$, $u^t_{i,succ(i)}$ can be computed successively with the equation $u^t_{i,succ(i)} = d^t_i + \sum_{j|succ(j)=i} u^t_{ji}$, starting from the leaf nodes in T_k.
- Set workload aggregation variable $s^t_{\tilde{i}} = \sum_{i \in T_k} x^k_i$; otherwise, $s^t_i = 0$ for all $i \in T_k, i \neq \tilde{i}$.

In a similar manner, Constraints (3-8) and (3-9) are automatically satisfied because they are equivalent to Constraints (3-3) and (3-4) in $\mathrm{P}_{\mathrm{STB}}$. The last two steps guarantee the satisfaction of Constraints (3-7) and (3-11). Constraint (3-10), which forces the flow to be unsplittable is guaranteed by the tree solution of $\mathrm{P}_{\mathrm{STB}}$.

It is thus proven that there is a solution to $\mathrm{P}_{\mathrm{NFB}}$ if and only if there is a solution to $\mathrm{P}_{\mathrm{STB}}$. Given the assumption of symmetric link cost and the objective function that minimizes the controller cost and link cost, at the optimality both formulations have the same objective function value. Therefore, the equivalency of the optimal

solutions of both formulations is proven.

3.4 Branch-and-Cut Algorithm for P_{STB}

The P_{STB} has an exponential number of cycle elimination constraints, and it is impractical to exhaustively enumerate all of them a priori. A branch-and-cut algorithm is then proposed for P_{STB}, which incorporates a dynamic constraint generation process into a typical branch-and-bound search of a MIP solver, e.g. Cplex, Xpress, etc.

By dropping the cycle elimination constraint, the relaxed version of P_{STB} can be solved with the branch-and-bound process. At each branching node of the branch-and-bound search tree, a separation routine is called to find the violated cycle elimination constraints and add them to the relaxed problem. The constraints generated are supposed to maintain the solution feasibility to the original P_{STB} and to tighten the lower bound, thus efficiently reducing the size of the branch-and-bound search tree.

The framework for solving P_{STB} is described below:

Branch-and-Cut Framework for P_{STB}
1. Solve the relaxed version of P_{STB}, i.e. by dropping the cycle elimination constraint, through the branch-and-bound search in the solver. 2. At each branching node of branch-and-bound process, run the separation routine and add violated constraints. The added constraints will be effective for the subsequent subproblems and cut off incumbent integer solutions that

are infeasible for the original formulation.

3. The procedure should stop after an optimal integer solution is proven, or after a maximum running time is reached.

In such a branch-and-cut algorithm, a separation problem is encountered at each branching node of the branch-and-bound tree as the violated constraints need to be identified. Although exact methods for such a problem exist, such as solving a max-flow min-cut problem (Lawler, 1985), a heuristic approach is still desirable in that exact methods are still computationally demanding.

The separation heuristic retrieves the solution at each branching node (not necessarily integer) and does the following tasks for each sector k :

- Construct a graph G' by using edges whose y_{ij}^k is greater than a pre-specified value.
- Identify all the connected components in G'.
- For each node in each connected component, find the smallest cycle Ψ, if exists, and then add the cycle elimination constraint, $\sum_{(i,j)\in E(\Psi)} y_{ij}^k \le |\Psi| - 1$, to the current problem.

Thus, to maintain the feasibility, when an integer solution is found at some branching node but is infeasible for the original formulation, the separation heuristic is called repeatedly at that node until no violation is found. The pseudo code for the separation heuristic is described below:

Separation Heuristic of Branch-and-Cut Algorithm for P_{STB}

Inputs: An intermediate solution (x, y, z) of $\mathrm{P}_{\mathrm{STB}}$ at a branching node; a threshold value γ, where $0 < \gamma < 1$.

Outputs: Violated cycle elimination constraints.

Steps:

For each $k \in \{1, ..., K\}$, do the following:

1. Construct a graph $G' = (N, E')$, where $E' = \{(i, j) \mid y_{ij}^k \geq \gamma\}$ and $N = \{i \mid 1, ..., I\}$.
2. $List := \{1, ..., I\}$.
3. If $List = \varnothing$, then exit. Otherwise, pick $i \in List$ and run a depth-first search at i on G' to find a set of nodes, C, such that all nodes in C are connected.
4. Update $List := List \setminus C$.
5. If there is no cycle in $G'' = (C, E'')$, where $E'' = \{(i, j) \mid i, j \in C \text{ and } (i, j) \in E'\}$, then go to Step 3.
6. For each $i \in C$, do the following:
   - Find a nontrivial path of fewest steps on G'' that starts and ends at i, if it exists. Denote as Ψ the set of nodes on such a path.
   - Generate and add to the problem the violated cycle elimination constraints: $\sum_{(i,j) \in E(\Psi)} y_{ij}^k \leq |\Psi| - 1$ for $k \in \{1, ..., K\}$. Go to Step 3.

3.5 Numerical Experimentation

3.5.1 Experiment Setup

To compare the performance of both formulations, we build a test network of 56 nodes, as shown in Fig. 3-4. Although the network of such a scale is not considered large for a typical network flow problem, it may be of an adequate size for our purpose since P_{FNB} considers multi-period demand patterns and can be seen as finding a best network structure and solving multiple network flow problems at once. Thus, the solution space of both formulations not only depends on the size of the underlying network but also on the number of periods under consideration.

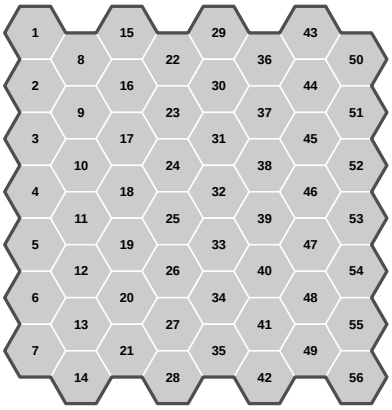


Figure 3-4 A Test Network of 56 Nodes

Demand profiles of five planning periods are randomly generated with a series of gamma distributions, which closely fit the real-world data shown later in Section 3.6. Specifically, the scale parameter of the gamma distribution gradually increases to generate the variation among the profiles, as illustrated in Fig. 3-5. From period 1 to 5, the mean and the variance of demand profiles increase. Descriptive statistics of each demand profile are summarized in Table 3-1.

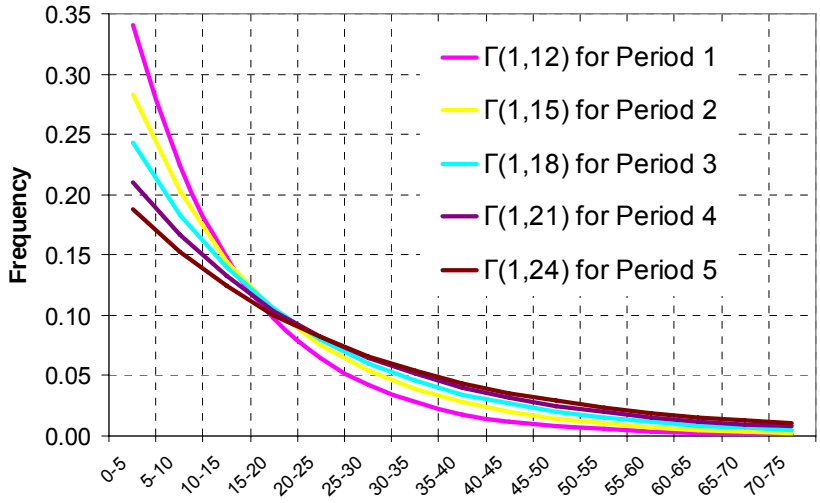


Figure 3-5 Distributions used for Generating Random Demand

Table 3-1 Descriptive Statistics of Drawn Demand Profiles

	Period 1	Period 2	Period 3	Period 4	Period 5
Count	56	56	56	56	56
Mean	11.5	14.7	19.3	22.6	24.9
Standard Deviation	9.47	14.37	20.09	21.90	24.10
Minimum	0.3	0.4	0.1	0.1	0.5
Maximum	41.0	63.6	88.4	94.3	102.3
Sum	645.6	823.8	1082.5	1264.6	1393.8

For each demand profile, 56 values are drawn, sorted in ascending order, and matched with the node indexes. This is done to create the variation within one profile, increasing demand intensity from up-left to bottom-right, which is visually expressed with a "heat map" in Fig. 3-6.

Link cost (or flow alignment penalty) is supposed to reflect traffic crossings between two adjacent nodes and to foster the tendency of combining two nodes with high traffic. In this randomly generated case, the value of crossing is set as the geometric

mean of demands at two end nodes of a link, i.e. $cross_{ij} = \sum_t \sqrt{d_i^t d_j^t}$, and the link cost parameter c_{ij} is calculated by the conversion function: $c_{ij} = 1 - cross_{ij} \Big/ \max\left\{cross_{ij} \text{ for all } (i,j) \in E\right\}$, which is monotonically decreasing with $cross_{ij}$ and translates the crossing information into a relative measure between 0 and 1.

For determining the tradeoff coefficient between two objectives, the rationale is provided as follows: If we assume F_1^U and F_2^U are the possibly minimum values for controller cost and flow alignment cost, respectively, we can call $\mathrm{F}^U = (F_1^U, F_2^U) \in \mathbb{Z}^2$ an ideal point for our multi-objective problem here. Such an ideal point is unattainable in general, so the next best thing is a solution that is as close as possible to this point (Marler and Arora, 2004).

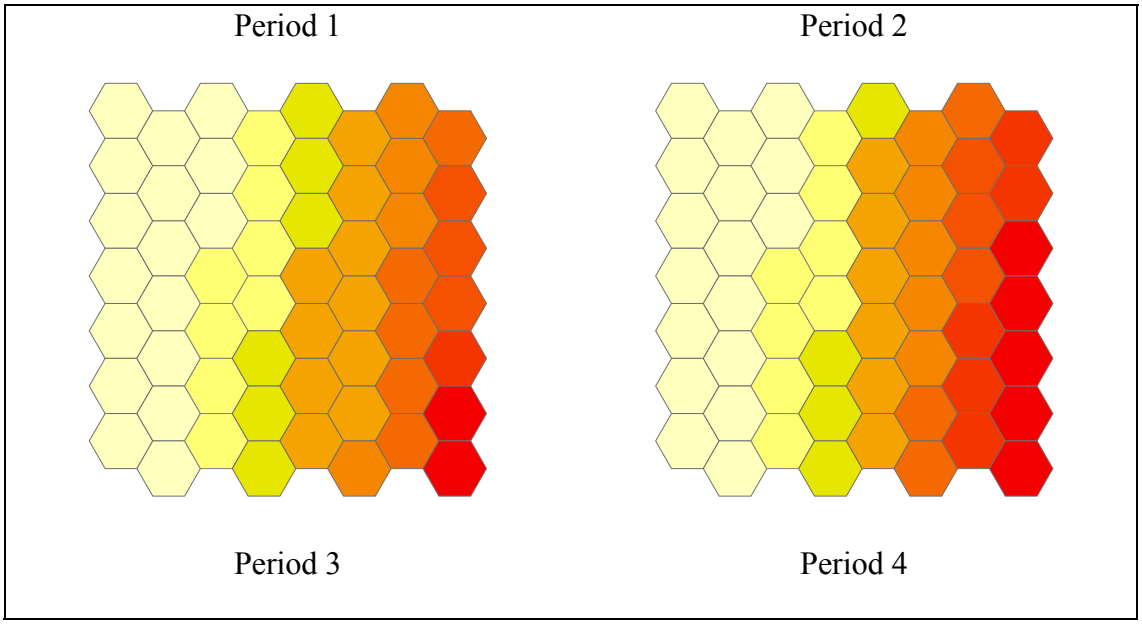
In addition, since two objectives in this study are measured in different units, a method for transforming into unitless metrics is needed. We take the transformation function from Koski and Silvennoinen (1987) for both objectives:

$F_i' = \dfrac{F_i}{F_i^U}$, where $F_i^U > 0$, i=1 for controller cost and i=2 for flow alignment cost.

Use of this transformation function implies that $\dfrac{F_2^U}{F_1^U}$ would be an adequate value for the tradeoff coefficient. Finally, in order to articulate a priori the preference that the objective one dominates two, we set $\mu = \dfrac{F_2^U}{F_1^U} C$, which uses a constant C to adjust

the weighing factor based on modeler's judgment. If F_1^U and F_2^U might not be known or provable, the currently best found values will be used as approximations.

For this experiment, at most two controller positions are considered per sector per period. In practice the capacity increase due to additional controllers has diminishing returns. To determine hypothetical values of sector capacity for facilitating this experiment, we take the sum of demand at mid-level (Period 3) divided by the expected number of sectors, which is three, and thus set at 330 the capacity value of one position. We also assume that the second position only adds 60% more capacity than the first one, so the capacity with two positions is 550 for this experiment.



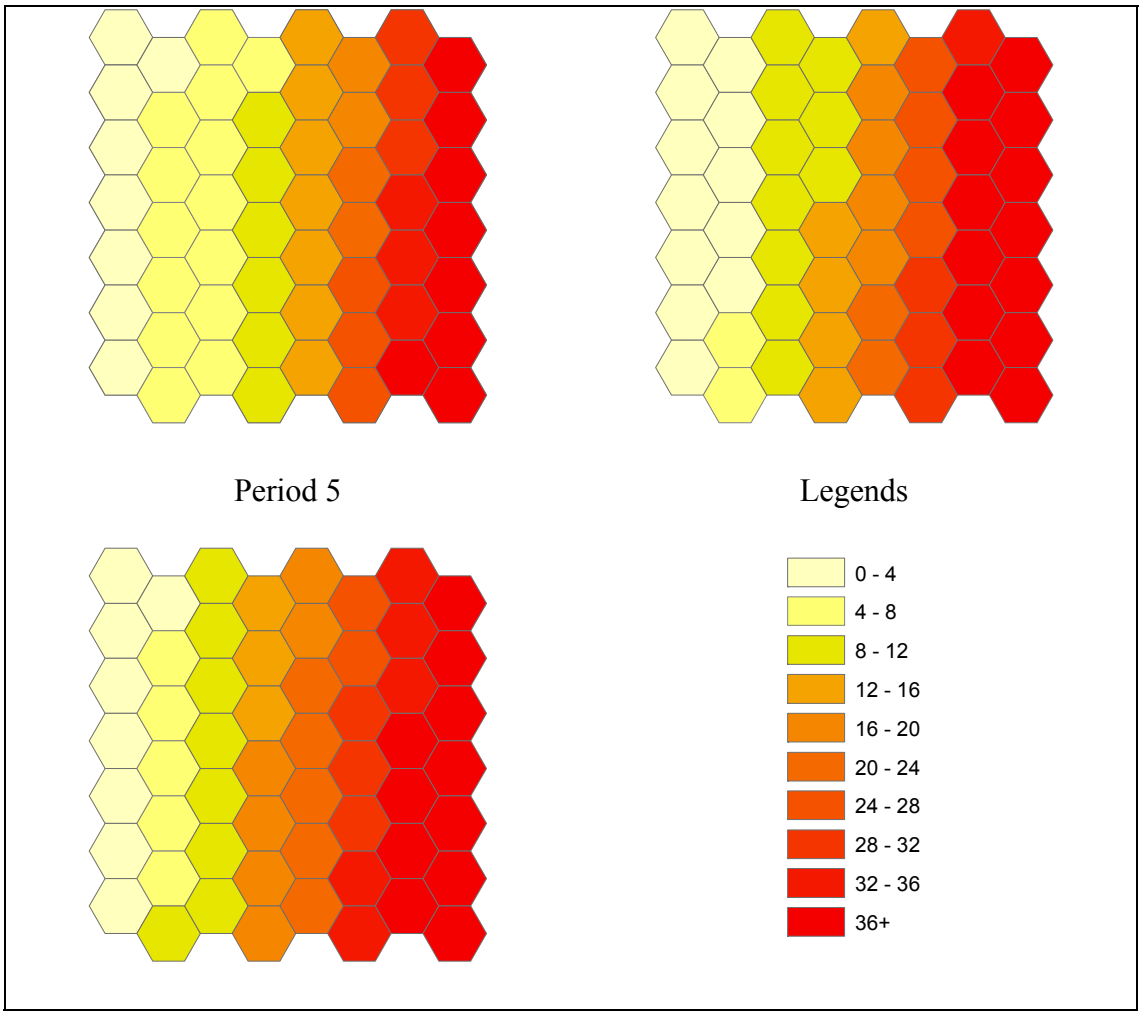


Figure 3-6 Heat Maps of Demand Profiles

3.5.2 Result Interpretation and Performance Comparison

Dealing with demand variation is one of the motivations of this study. Three cases are considered here: 1) high demand variation; 2) low demand; 3) moderate demand. Each of the cases is built by selecting 3 demand profiles in Table 3-1 and solving them with Xpress-Mosel ver. 2.4.1 on a Dell PowerEdge 1900 with Intel Xeon 2.66Ghz processor and 12 GB memory. (One core is used, and memory use never exceeds 1 GB.)

Fig. 3-7 illustrates the optimization result. Case 1 is intended to demonstrate how the model deals with demand variation, so it is designed by selecting time periods with

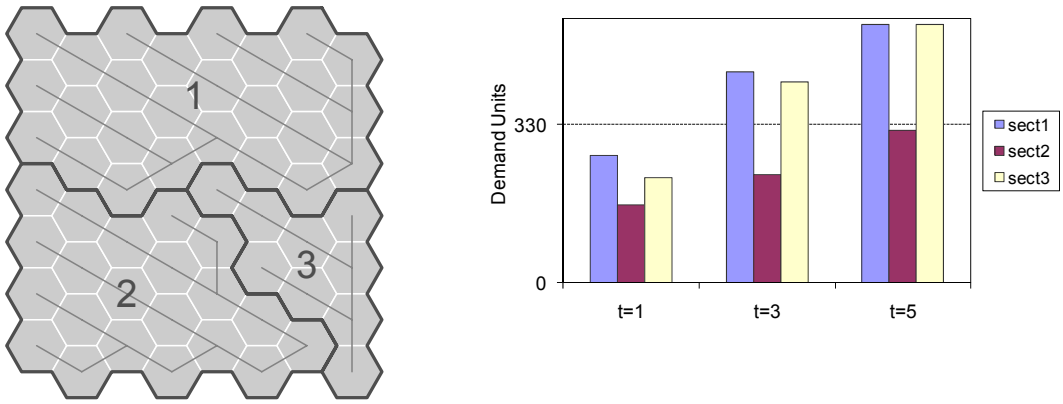
low, medium, and high demands. Three sectors are needed in this case. Sectors 1 and 3 use two controllers at t=3 and t=5 (see the sector demand above the dashed line, i.e. the capacity offered by one position).

To address a low demand situation, Case 2 is designed by selecting time periods {1, 2, 3}. It is natural to expect that the number of optimal sectors is below that in Case 1. Two-position sectors are used for certain time periods (t=2 and t=3).

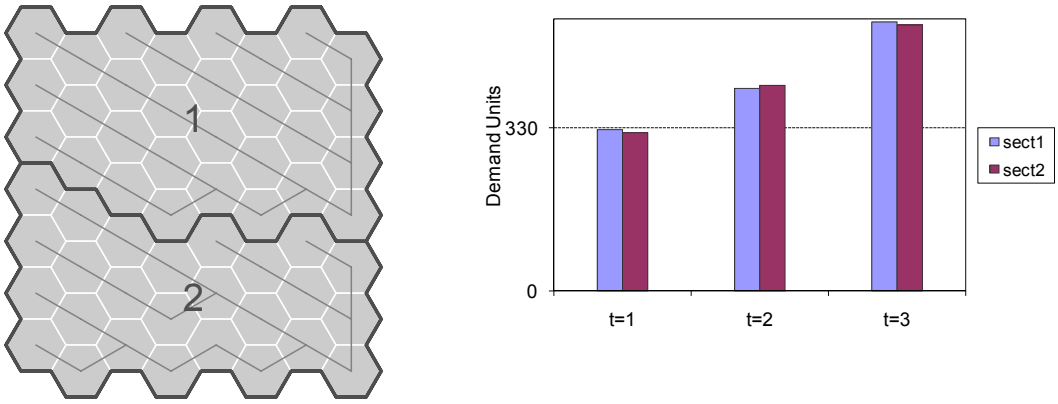
Case 3 is intended to address a moderate demand situation, where time periods {2, 3, 4} are selected. The number of optimal sectors is four, and all the sectors use one controller in each period. This seems to be counter-intuitive to our motivation since we assert that multiple positions increase capacity and might help reduce the number of sectors, thereby improving resource allocation. In fact, if the demand is relatively steady, using only one-controller sectors turns out to be a more efficient design. It reaffirms our hypothesis that the staffing strategy of multiple controllers per sector is only a mechanism for dealing with temporary demand peaks.

For the objective of flow alignment, the gray lines on the figures depict the solutions of link variables. As the demand increases from the top left to the bottom right, the shape of sectors generally captures this orientation.

(a) Case 1 – Time Period={1,3,5}



(b) Case 2 – Time Period ={1,2,3}



(c) Case 3 – Time Period ={2,3,4}

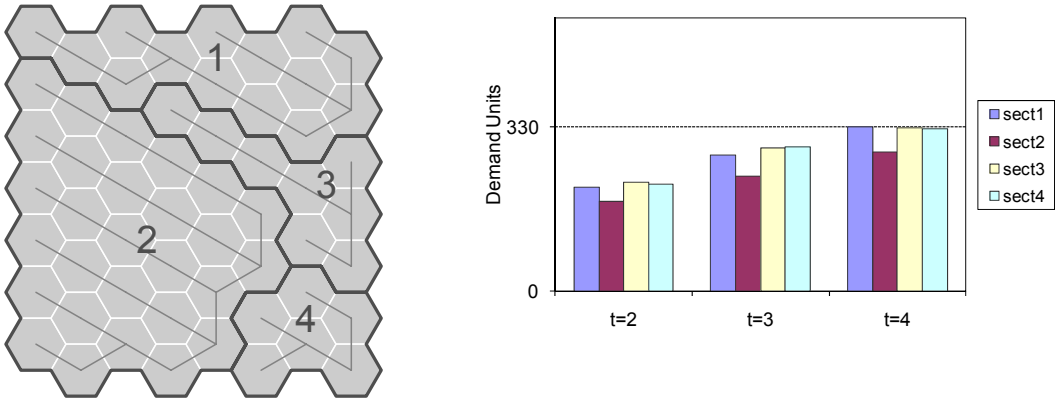


Figure 3-7 Summary for Three Cases of Demand Variation

In addition, the computation performance of both formulations is also of interest. Each formulation can be further enhanced to improve performance, i.e. by adding valid inequalities for P_{STB}, and reducing problem size for P_{NFB}.

When solving P_{STB}, it is observed that prioritizing certain variables during the B&B search yields a short convergence time. The best strategy found for selecting variables is to first branch on the controller variables $z_p^{k,t}$ followed by the link variables y_{ij}^k. The controller variables are more important in helping the search due to the following reasons:

- Controller staffing determines the sector capacity;
- Controller cost dominates the flow cost in the objective function.

Identifying valid inequalities could also help improve the performance of the B&B algorithm. If the valid inequalities are well-chosen, then the LP bound should be improved, and the B&B algorithm is more effective by pruning the nodes to be visited. (Wolsey, 1998) We have identified several valid inequalities for P_{STB} to accelerate the convergence time of the B&B algorithm:

- $\sum_k y_{ij}^k \le 1, \forall (i,j) \in E$: It limits at most edge (i,j) will be assigned one sector. Although this condition is implicitly ensured by other constraints in P_{STB}, adding this valid inequality improves the bound of the LP relaxation.
- $x_i^k \le z_{p=1}^{k,t=1}, \forall i,k$: It says that if sector k is not chosen to be open, then sector k will not receive any assignment from node i.

- $z_{p=1}^{k,t=1} = z_{p=1}^{k,t+1}$, $\forall k,\ t = 1,...,T-1$: It seems trivial but helps the branching of the controller variables. The formulation implies once a sector is open, it will be open for all the time periods.

When solving $\mathrm{P_{NFB}}$, the fact that all the nodes can be sink nodes fosters the occurrence of the solution symmetry, which means many solutions yield the same objective function value and impacts the performance of the B&B algorithm. To reduce solution symmetry as well as to reduce problem size, we can pre-specify a subset of nodes S, called "seeds", which represents a set of candidate sink locations to be determined by the optimization model, and solve the modified $\mathrm{P_{NFB}}$.

Modified $\mathrm{P_{NFB}}$:

$$\min f(u,s,w,g) = \mu \sum_{i\in S,p,t} h_p g_{ip}^t + \sum_{(i,j)\in E} c_{ij} w_{ij}$$

Subject to:

$$\sum_{j\in\delta(i)} u_{ji}^t + d_i^t - \sum_{j\in\delta(i)} u_{ij}^t = \begin{cases} s_i^t & \text{for all } i\in S, t\in\{1,...,T\} \\ 0 & \text{for all } i\in\{1,...,I\}\setminus S, t\in\{1,...,T\} \end{cases} \tag{3-12}$$

$$g_{i,p+1}^t \le g_{i,p}^t \quad \text{for all } i\in S,\ p\in\{1,...,P-1\},\ t\in\{1,...,T\} \tag{3-13}$$

$$s_i^t \le \sum_{p=1} U_p g_{i,p}^t \quad \text{for all } i\in S,\ t\in\{1,...,T\} \tag{3-14}$$

$$\sum_{j\in A_i} w_{ij} = \begin{cases} 1 - g_{i,p=1}^{t=1} & \text{for all } i\in S \\ 0 & \text{for all } i\in\{1,...,I\}\setminus S \end{cases} \tag{3-15}$$

$$u_{ij}^t \le \left(\sum_p U_p\right) w_{ij} \quad \text{for all } i\in\{1,...,I\},\ j\in\delta(i), t\in\{1,...,T\} \tag{3-16}$$

This seed and demand-flow technique was first introduced by Yousefi et al. (2007) to maintain the connectivity of the hex-cells grouped to a sector. However, their model is very sensitive to the seed locations because, as mentioned in Section 3.3, its objective function capitalizes the flow carried on the links. On the other hand, P_{NFB} is relatively indifferent to the link flow, as it only considers a fixed cost if a link is used. Thus, as long as the seed locations are closely related to the demand distribution over the planning area, this "seed" technique does not prevent us from finding the optimal solution to the original P_{NFB} formulation.

We randomly generate 10 instances by using the demand distributions in Fig. 3-5 for three test cases. The results of computational experiments are summarized in Table 3-2. By observing the average performance in Fig. 3-8, we find that one formulation is no better than the other. In some instances, P_{STB} converges faster than P_{NFB}, but the computation time of P_{STB} has a wider range.

Table 3-2 Computational Performance of Two Formulations

Case		t={1,2,3}		t={1,3,5}		t={2,3,4}	
		STB	NFB	STB	NFB	STB	NFB
Parameter Setting	No. of Seeds for NFB or No. of Candidates for STB	4	4	4	4	6	6
Computation Time* (Seconds)	Mean	313.56	260.31	156.02	330.82	1,828.20	756.10
	Min	27.93	129.52	54.28	150.08	127.34	502.12
	Max	450.23	373.44	250.16	861.20	5,202.97	1,428.38

* Convergence time at 5% MIP gap; Valid inequalities employed for STB; Seed strategy adopted for NFB; Average over 10 randomly generated cases.

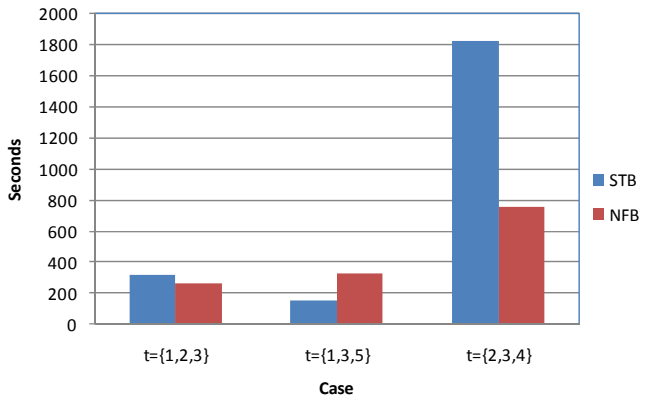
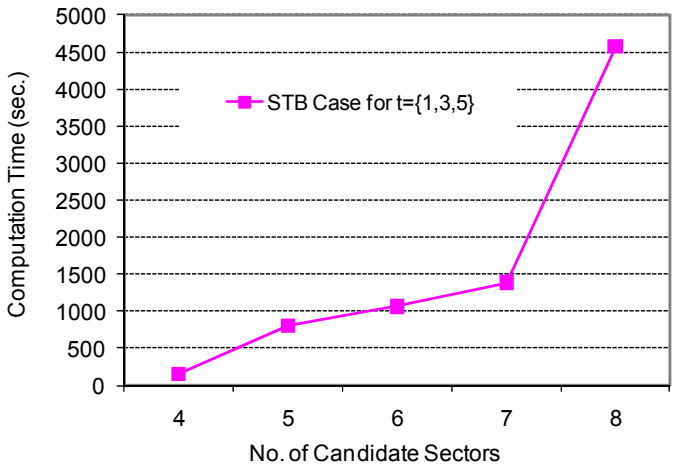


Figure 3-8 Average Computation Time for Three Cases

The size of P_{STB} grows with the number of candidate sectors. The case of t={1,3,5} is used for sensitivity study. Again, we randomly generate 10 instances and vary the number of candidate sectors from 4 to 8, where the optimal number of sectors is known as 3. In Fig. 3-9, the convergence time increases more than linearly with the number of candidate sectors, which is due to the solution symmetry. That is, assigning a node to either Sector A or Sector B, for example, yields the same objective function value. This symmetry of solutions does not help the branch-and-bound algorithm converge. Thus, unwisely choosing this parameter for P_{STB} would impact the computational performance.



(Convergence time at 5% MIP gap; Valid inequalities are employed for STB. Average of 10 randomly generated cases of {1,3,5}.)

Figure 3-9 Sensitivity Analysis of P_{STB} Performance vs. Number of Candidate Sectors

Another aspect of the models we should examine is the number of time periods (or demand profiles) since one of the research objectives in this study is to consider multi-period demand patterns. Fig. 3-10 summarizes the computation times of both formulations as the number of time periods increases. Overall, P_{STB} maintains a relatively steady trend, as opposed to P_{NFB}. When considering an additional period, P_{STB} only requires one additional constraint, i.e. constraint (3-4) and a set of controller variables. On the other hand, P_{NFB} requires the addition of almost all the constraints and variables associated with the time index, so its formulation expands linearly with the number of periods under consideration.

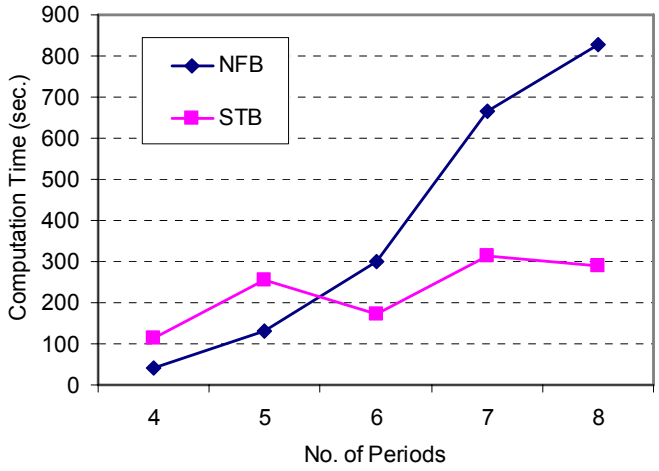


Figure 3-10 Sensitivity Analysis on the Number of Time Periods

3.5.3 Discussion

We have analyzed the theoretical properties of both formulations and conducted computation experiments to understand their performance. The P_{STB} assigns nodes and links to individual sectors while maintaining connectivity of each sector, so its size grows rapidly with the potential number of sectors. However, its property of explicitly identifying node-sector and link-sector assignments in the formulation, unlike The P_{NFB} requiring post-processing effort, would be beneficial to some design criteria that require directly dealing the physical location of the nodes, links or even sectors, e.g. maintain sector aspect ratio, counting flights that cross adjacent sectors.

The P_{NFB} considers demand (workload) units at each node as flow commodities in a network flow problem. Unlike P_{STB} which needs an ad hoc process for dynamically generating the cycle elimination constraints, P_{NFB} can be directly implemented in any integer programming solver without extra modification efforts. Since commercial solvers utilize many successful heuristics in generating valid inequalities and in

exploring the branch-and-bound tree, our experience shows that the solver always finds the first few feasible solutions faster for P_{NFB} than for P_{STB} which is solved with the branch-and-cut algorithm. Nevertheless, the convergence performance depends on the instances and is inconclusive for either formulation.

For P_{NFB}, the formulation size as well as the computation time grows rapidly with the number of planning periods. The formulation size of P_{STB} is relatively insensitive to the number of periods, but it grows rapidly with the potential number of sectors.

Note that it is straightforward for P_{STB} to identify node-sector assignment because of the formulation. For P_{NFB}, by labeling the flow variables of each link with the origin information, we can also turn P_{NFB} into a multi-commodity network flow problem and see where the commodities come and go. This treatment would have the disadvantages of using at least $I \times |E|$ variables and associated flow conservation constraints to identify the origin of flow commodities on each link and would certainly complicate the problem by expanding its size.

3.6 Real-World Application

3.6.1 Experiment Setups

To bring a practical sense to real-world application, we test our models on a realistic problem size. One of the proposed MPVC models is implemented on historical traffic data (1-minute radar positions) recorded in the Washington DC enroute Center (ZDC) on April 21, 2005 (in GMT). The control area of ZDC is tiled with 1043 hex cells of equal size, 41 of which are selected as seeds and evenly distributed within the design

area, as in Fig. 3-11. The demand at each cell is measured as the number of TZ radar hits between FL240 and FL360. For a given time period, the number of hits in a cell or a sector implies not only aircraft counts but also aircraft dwell time. To reduce problem complexity, we use aircraft position hits as a surrogate for workload. Yousefi (2005) has shown that, for small airspace cells aircraft count is highly correlated to a composite workload measurement by more elaborate traffic complexity metrics. (In fact, before the airspace cells are clustered into sectors, there are very few options for complexity metrics.) For model demonstration, we used the number of radar hits as a surrogate for measuring sector demand and capacity.

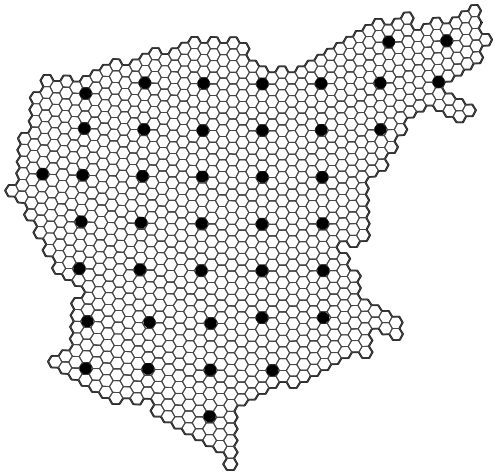


Figure 3-11 ZDC (Without Ocean) and Seed Locations

Fig. 3-12 illustrates the temporal demand magnitude at ZDC every 2 hours. This is a series of histograms (each one running toward the reader) for various times of the day. Each histogram gives the frequency of radar hits. Fig. 3-13 shows the variation in traffic patterns (mainly intensity) in ZDC on that day.

In this experiment, the formulation P_{NFB} is used with a slight modification of a cost term in the objective function. In a later section, we will compare the performance of

our model to an existing one proposed by Yousefi et al. (2007), so we apply their definition of flow alignment penalty in P_{NFB} to create a fair comparison basis. Specifically, the modified objective function capitalizes the flows of the links u_{ij}^t, instead of a fixed cost, and is written as:

$$\min f(u,s,u,g) = \mu \sum_{i \in S,p,t} h_p g_{ip}^t + \sum_{(i,j) \in E,t} c_{ij}^t u_{ij}^t$$

Two demand data sets were created to show how MPVC performs when demand variation is high or low.

To demonstrate the multi-controller effect, at most two controller positions could be used in each resulting sector, i.e. $p \in \{1,2\}$, so there are two sets of capacity choices for each seed identified in Fig. 3-11.

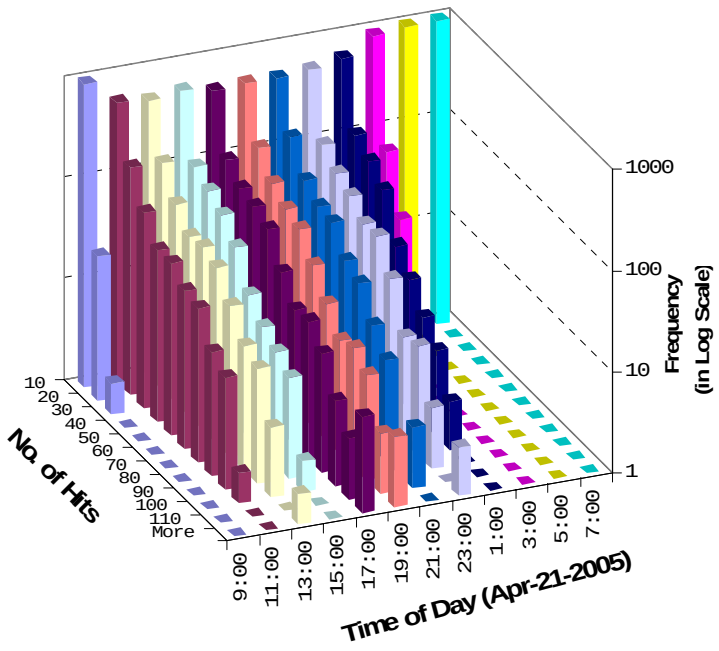


Figure 3-12 Period-Wise Histograms of Radar Hits at ZDC

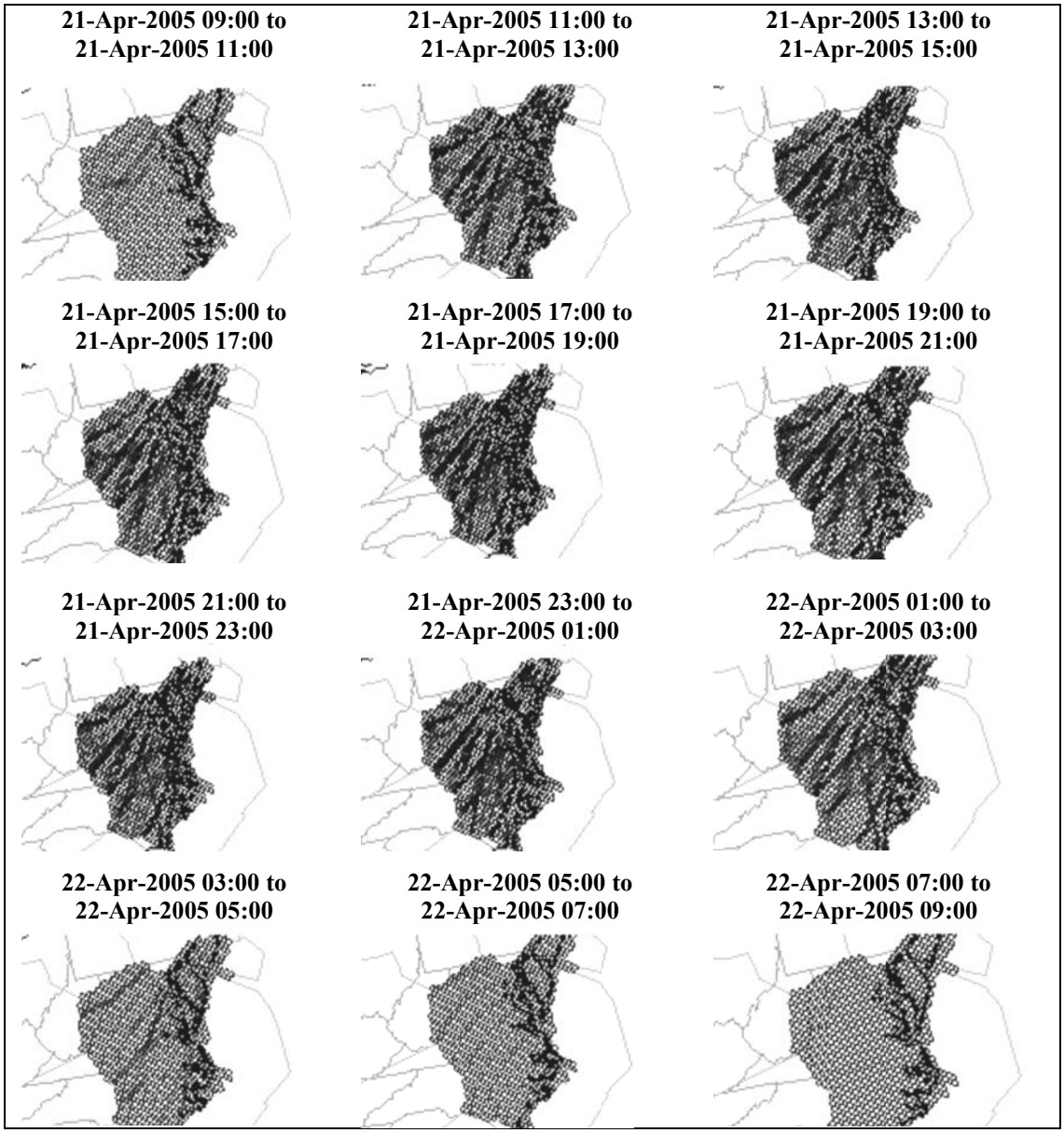


Figure 3-13 ZDC Traffic Patterns on April 21, 2005 (in GMT)

3.6.2 High-Demand Variation Case

For this case, we divided the 16-hour busy period of the day (11:00 to 03:00 the next day) into four 4-hour intervals. The controller unit costs in the objective function are set as $h_1 = 1$ and $h_2 = 0.9 \le h_1$. As mentioned earlier, these coefficients should reflect the cost or number of controllers. The assumption that a 2-controller team is less

costly than two 1-controller teams is not necessary, but helpful for solving the integer program quickly. If a temporary capacity increase is needed, then the tradeoff between cost and capacity gained will be determined through the optimization process.

Ideally, the sector capacity values should be estimated with controller capability. (An estimation model for sector capacity will be proposed in a later chapter.) Given the 4-hour period length, precise estimates for sector capacity might be impractical and under large variation. To facilitate model demonstration, the capacity values from the current operational environment are approximated as follows. ZDC has about 17 enroute sectors between FL240 and FL360. (This varies with time of day, and not all enroute sectors lie in this altitude range.) For the capacity provided by 1-controller team, we pick the highest total demand amongst the design periods and divide this number into total demand for each period:

U_1=max {total demand in period t} / 17

This is a very conservative estimate since it implicitly assumes that at the busiest period one controller serves one sector in average. The capacity of a 2-controller team should incorporate the diminishing effect on productivity of an additional controller. In this study, we assume an additional controller increments capacity by 60%, i.e.

$$U_2 = U_1 \times 0.60$$

To let sector shapes align with air traffic flow and let the model connect cells with high aircraft transfer, the cost of link between two cells at each period, c_{ij}^t, is defined as the inverse of the total number of aircraft crossings from both directions within the defined period. This setting is suggested by Drew (2008) and will improve sector

shape without impacting the design objective. The cost tradeoff parameter μ is set to a high value 10^6, which allows controller cost to dominate the objective, flow alignment.

We solved this instance of MPVC with Xpress-Mosel solver software on a Dell PowerEdge 1900 with Intel Xeon 2.66Ghz processor and 12 GB memory. Solver time was 45,578 seconds (12.6 hours) for a 12.17% optimality gap. Table 3-3 summarizes the controller requirements. The resulting sectors are shown in Fig. 3-14, and demand distribution is displayed in Fig. 3-15 with dashed lines depicting the assumed capacity values.

Three of the 17 sectors formed in the optimization employed 2-controller teams. The total number of controller hours used was (20+19+20+18)×4 = 308.

Table 3-3 MPVC Controller Requirements for High-Demand Variation Case

Resulting No. of Sectors	Resulting No. of Controller Shifts				Capacity Assumed	
	11:00 | 15:00	15:00 | 19:00	19:00 | 23:00	23:00 | 03:00	Using 1 Position	Using 2 Positions
17	20	19	20	18	2315	3704

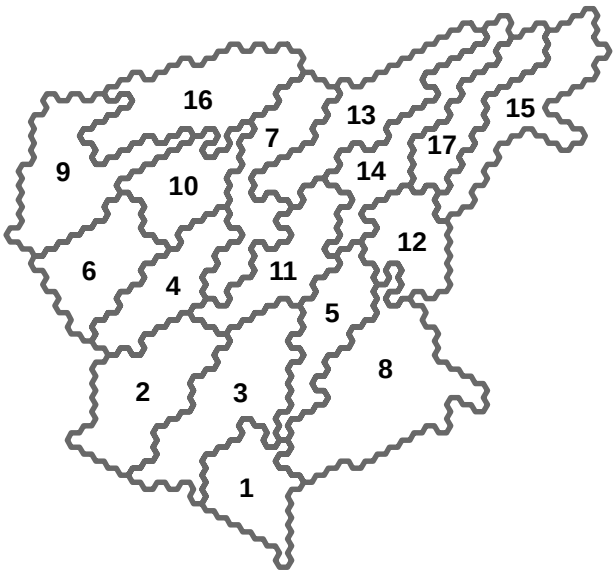


Figure 3-14 MPVC Sector Boundaries for High-Demand Variation

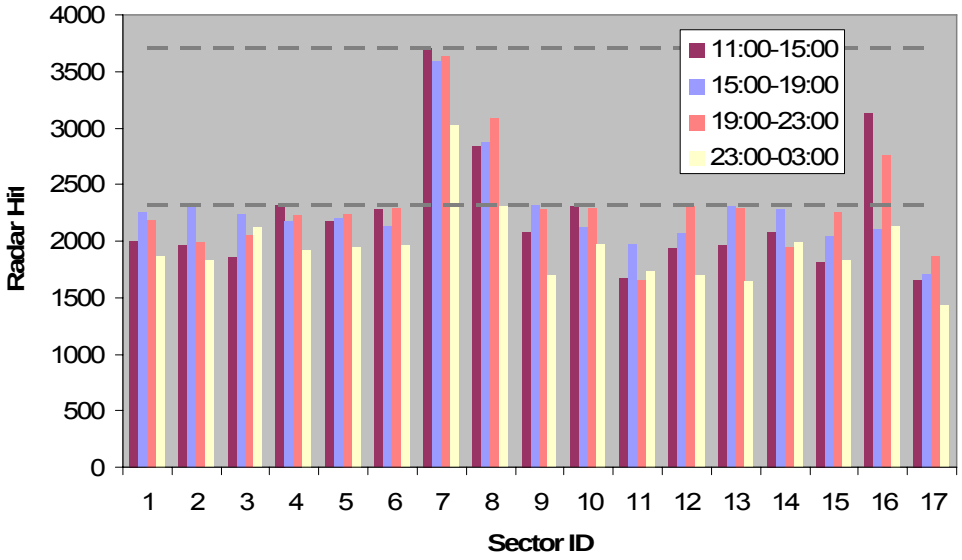


Figure 3-15 MPVC Sector Demand Distribution for High-Demand Variation

3.6.3 Low-Demand Variation Case

For this case, we divided the low-demand period of the day (17:00 to 01:00 the next day) into four 2-hour time intervals. This was constructed to check a hypothesis that

when demand is steady (low variation), 1-controller teams make efficient use of controller resources.

MPVC was run with the same settings as the high-demand case. The results in Table 3-4 suggest that the optimized number of sectors was 18. Note that this low-variation case created one more sector than the high-variation case (18 vs. 17). When demand is steady but high, creating two 1-controller sectors is more efficient than one 2-controller sector because the two sectors have greater capacity. (Recall that the second controller adds marginally less capacity than the first controller.) This principle, demonstrated by this proposed model, captures the current practice of splitting sectors during busy periods.

The resulting sectors are depicted in Fig. 3-16 and demand distribution is displayed in Fig. 3-17. During this 8-hour planning horizon, the total number of controller hours required was (19+18+18+18) ×2 = 146. Only one sector during one period required a 2-controller team. This confirms our hypothesis that one controller per sector suffices when the traffic is busy but less variable across the planning horizon.

Table 3-4 MPVC Controller Requirements for Low-Demand Variation Case

Resulting No. of Sectors	Resulting No. of Controller Shifts				Capacity Assumed	
	17:00 | 19:00	19:00 | 21:00	21:00 | 23:00	23:00 | 01:00	Using 1 Position	Using 2 Positions
18	19	18	18	18	1272	2035

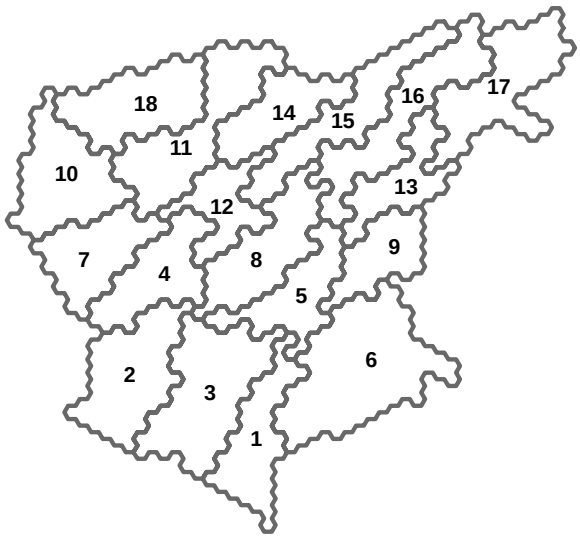


Figure 3-16 MPVC Sector Boundaries for Low-Demand Variation

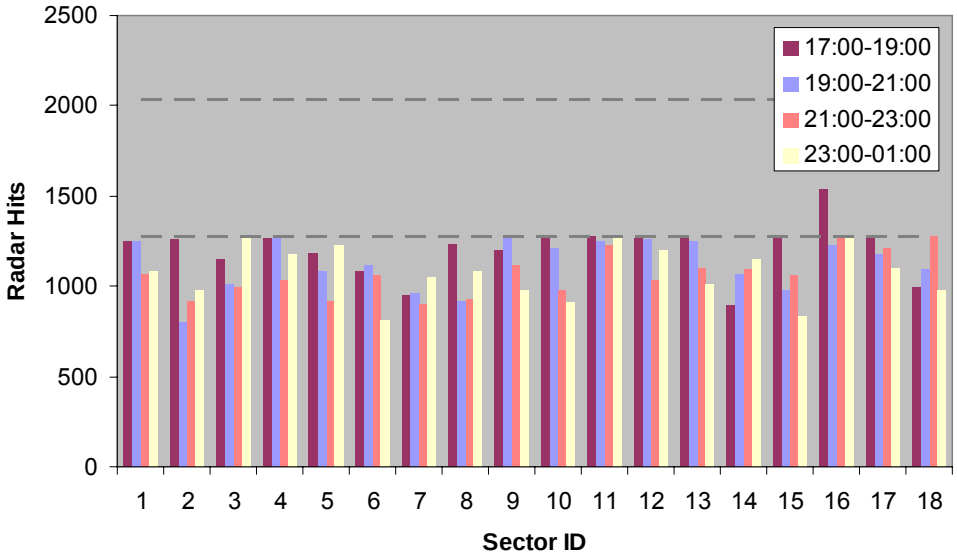


Figure 3-17 MPVC Sector Demand Distribution for Low-Demand Variation

3.6.4 Comparison with Alternate Design Concept

To compare our MPVC results with a policy that balances average workload across the 16 hours, we altered the modified P_{NFB} and ran the mixed integer program (YMIP) of Youefi et al. (2007) on the same data set with aggregated demand. Neither

multi-period demand nor controller staffing are considered, so we set $T = 1$ and $P = 1$

.

YMIP accepts the number of sectors (17, in this case) as input. This is enforced by Constraint (3-17). The primary objective of YMIP is to align sectors with flows. Workload balance is a secondary objective in YMIP, addressed in the constraints as maximum deviation from the average workload across all sectors $\mathrm{W}_{\text{target}}$, so the Constraint (3-14) is replaced by Constraint (3-18). The tolerance parameter for workload balancing is set to $\gamma = 0.05$, i.e. total workload in each sector for the planning horizon will be within 5% of the average over all sectors.

$$\sum_{i \in S} g_{i,p=1}^{t} = \text{ Desired No. of Sectors} \tag{3-17}$$

$$g_{i,p=1}^{t} \cdot (1-\gamma) \cdot \mathrm{W}_{\text{target}} \le s_i^t \le g_{i,p=1}^{t} \cdot (1+\gamma) \cdot \mathrm{W}_{\text{target}} \quad \text{for all } i \in S \tag{3-18}$$

Fig. 3-18 shows resulting sectors from YMIP, while the demand distribution is shown in Fig. 3-19. Table 3-5 shows that (24+24+26+17)×4 = 364 controller-hours are required to serve the demand over the planning horizon. That is 56 more (worse) than the 308 required by MPVC. The less efficient use of controller-hours can be attributable to the unacknowledged demand variation over time.

Table 3-5 YMIP Controller Requirements for High-Demand Variation Case

Resulting No. of Sectors	Resulting No. of Controller Shifts				Capacity Assumed	
	11:00 | 15:00	15:00 | 19:00	19:00 | 23:00	23:00 | 03:00	Using 1 Position	Using 2 Positions
17	24	24	26	17	2315	3704

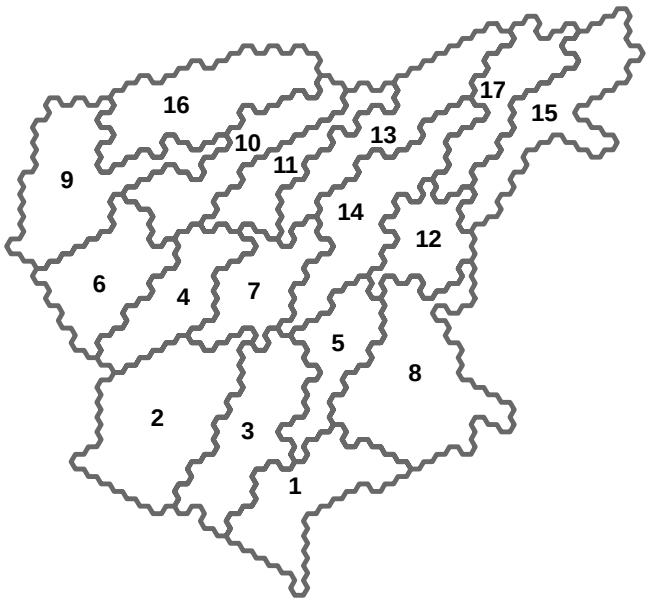


Figure 3-18 YMIP Sector Boundaries for High-Demand Variation

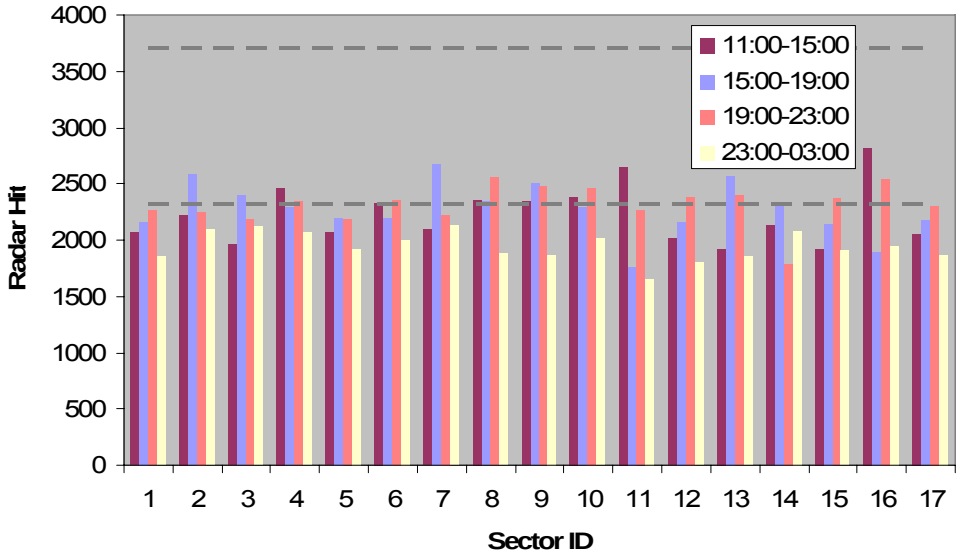


Figure 3-19 YMIP Sector Demand Distribution for High-Demand Variation

Note that an alternate sector design strategy using YMIP would be to increase the number of sectors so that no sector workload level will exceed the 1-controller threshold of 2315, which is used in the previous MPVC example. We tested this and found that 20 sectors would be required (each with a one controller) for a total of 320

controller hours. This is better than the 364 controller-hours under the 17-sector YMIP policy, but still higher (worse) than the 308 achieved by MPVC under a multi-controller policy.

For the low-demand variation case, again for comparison purposes, we ran the YMIP model, but this time calling for 18 sectors. Table 3-6 shows the controller requirements per time period. This sectorization requires (23+21+19+18)×2 = 162 controller-hours to serve the demand in the planning horizon. This is 16 more controller-hours than MPVC required. The YMIP sectors are shown in Fig. 3-20. Fig. 3-21 shows the demand distribution over time. Each bar over the lower dashed line indicates need for a 2-controller team.

Table 3-6 YMIP Controller Requirements for Low-Demand Variation Case

Resulting No. of Sectors	Resulting No. of Controller Shifts				Capacity Assumed	
	17:00 | 19:00	19:00 | 21:00	21:00 | 23:00	23:00 | 01:00	Using 1 Position	Using 2 Positions
18	23	21	19	18	1272	2035

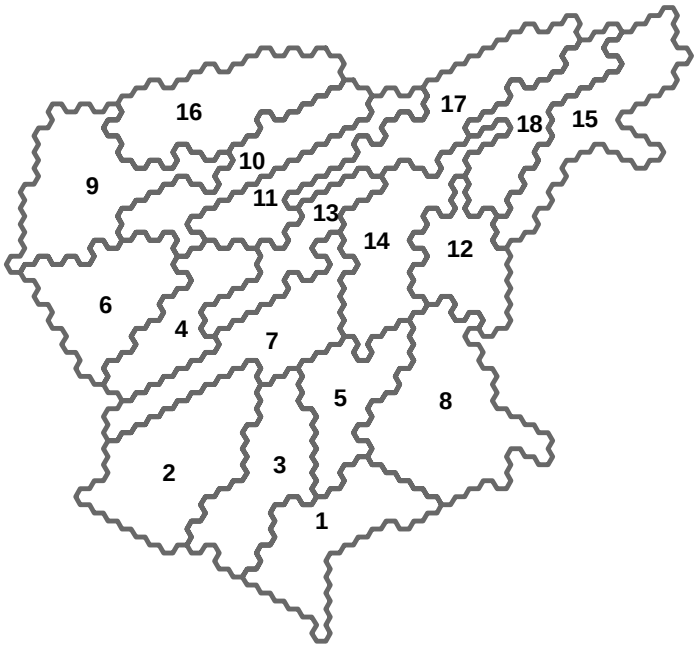


Figure 3-20 YMIP Sector Boundaries for Low-Demand Variation

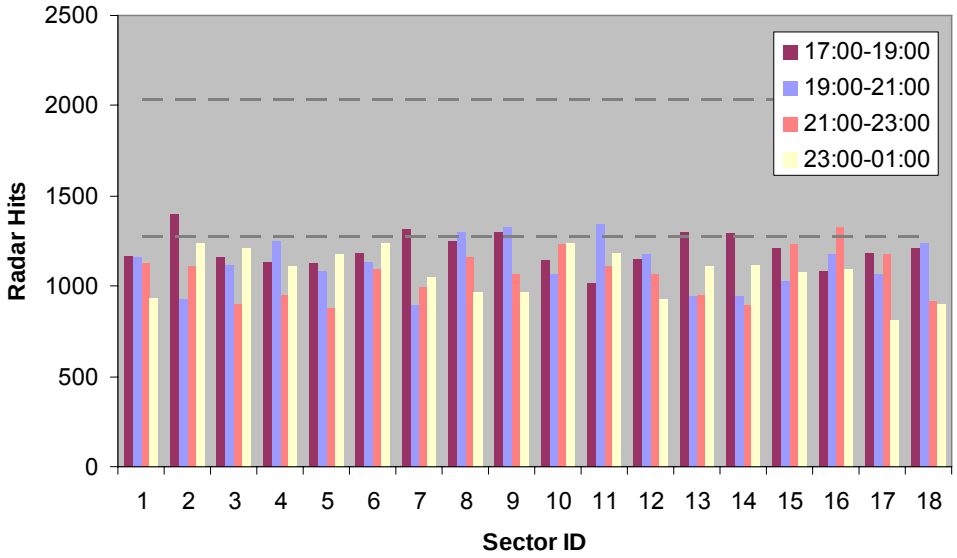


Figure 3-21 YMIP Sector Demand Distribution for Low-Demand Variation

Table 3-8 summarizes the numerical results. The primary statistic is the number of controller-hours. The two optimal (minimal) values for the two test cases are highlighted. In addition, we show the average aircraft dwell time by using whole day

traffic. For each sector, we compute the average dwell time of the aircraft trajectories, then average these over all sectors. This is a simple but reasonable surrogate for flow alignment (alignment with traffic flows tends to increase sector dwell time). The two models are comparable in flow alignment, a common objective of the two models. For YMIP, this is traded off with workload balancing; for MPVC, this is traded off with controller cost. The tradeoff in each model can be controlled by parameter settings.

The balance deviations are shown in the last two rows of Table 3-7. These are, respectively, the maximum (positive) deviation and minimum (negative) deviation from average workload (radar hits) computed across all sectors over the planning horizon. This is simply confirmation that YMIP has balanced workload to within its reasonable tolerance, but that MPVC has deliberately unbalanced sectors to allow for larger sectors that require multiple controllers.

For the above experiments, it has been shown that, given the time-varying nature of traffic, the sectorizations from the proposed model not only accommodate the multi-period demand but also consider the overall efficiency of controller staffing requirements. An aggregated model, such as YMIP, neglects demand variation and might produce an inefficient design in terms of controller-hours. In the case with low-demand variation, when designing sectors from a clean sheet, one controller per sector is a more effective choice than applying multiple controllers.

Table 3-7 Summary of Numerical Results

Test Case	High Demand Variation		Low Demand Variation	
Planning Horizon	16 Hrs		8 Hrs	
Duration per Period	4 Hrs		2 Hrs	
Model (MIP)	MPVC	YMIP	MPVC	YMIP
Design Objective	Minimize no. of controller shifts and sectors; Minimize flow alignment cost	Balance workload among sectors; Minimize flow alignment cost	Minimize no. of controller shifts and sectors; Minimize flow alignment cost	Balance workload among sectors; Minimize flow alignment cost
Required Controller-hours	308	364	146	162
Avg. Flight Dwell Time	8.0	8.5	7. 8	8.2
BalDev+	59.1%	5.0%	18.8%	5.0%
BalDev-	-23.7%	-5.0%	-13.4%	-5.0%

3.7 Summary and Contributions

Demand variation and controller staffing are closely related, but this relation is rarely addressed in the existing literature on airspace design. In this chapter, the problem of sectorization with consideration of multi-period demand patterns and time-varying controller staffing is defined. Optimization models are proposed for a clean-sheet airspace sectorization to take into account demand patterns period by period and to find an efficient controller staffing plan to accommodate traffic variation.

Under the proposed sectorization approaches, the controller costs are minimized, which should be of great interest to air navigation service providers. A secondary design objective is also addressed by grouping hex-cells into contiguous sectors with shapes that tend to align with traffic flows, thereby increasing sector dwell time and

minimizing controller handoffs. The proposed approaches avoid frequent and disruptive wholesale resectorization. Sector boundaries can remain in place throughout the day, avoiding drastic boundary changes over time. This design concept differs from the workload-balancing sectorization in the literature by capitalizing on the fact that sector capacity varies with the discrete number of controllers working that airspace.

The main contributions from this chapter can be summarized as follows:

- The problem of sectorization with multi-period demand patterns and variable sector capacity choices (MPVC) is formally defined, and its complexity is proven.
- Two integer program formulations P_{STB} and P_{NFB} are proposed to address the defined MPVC. P_{STB} is based upon maintaining a spanning tree for each resulting sector whose capacity is to be optimized while P_{NFB} expands the typical network design problem by adding dummy nodes and links to represent potential sector locations and capacity values. Both formulations solve MPVC, and their equivalency is proven.
- A solution technique is developed for P_{STB} that avoids the need to enumerate an exponential number of cycle elimination constraints for P_{STB}. Specifically, the relaxed version of P_{STB} is solved with the branch-and-cut algorithm that dynamically generates violated constraints by using the developed separation heuristic, so the feasibility of the original P_{STB} can be maintained.

- The performance of both formulations is examined with randomly generated numerical examples. While P_{STB} is insensitive to the number of periods (or demand patterns), which is a good property for a multi-period design, P_{NFB} does not increase its size exponentially and is more suitable for handling a large scale network.
- Since each formulation has its advantages, the choice of formulations depends on application areas. For example, P_{NFB} does not require dynamic constraint generation, so its application to a realistic size problem is still computationally tractable. P_{STB} explicitly addresses the assignment of nodes and links in the formulation, so it is applicable to problems that require such information during optimization, such as the sector combination problem to be introduced in a later chapter.
- The proposed design objectives (controller cost minimization and flow alignment) are confirmed on real traffic data from Washington Center. Specifically, we compared performance with a sectorization strategy that does not take demand variation into account. The numerical experiment using assumed controller capability values has demonstrated that the resulting sectorization created a design comparable to those of competing models in terms of flight dwell time and flow alignment, but saved 10%–16% controller-hours, depending on the degree of demand variation over time.

Chapter 4: Heuristic Based on Mathematical Programming for MPVC

This chapter is organized as follows: Section 4.1 describes the need for a heuristic approach in generating quality solutions within time constraints. Section 4.2 proposes a heuristic based on mathematical programming that involves a large-neighborhood search. This heuristic is then applied in a parallel computation framework. Some key components of the heuristics that will impact the performance are discussed. In Sections 4.3, 4.4, and 4.5, various neighborhood definitions are proposed and their performances in improving upon a starting solution are compared. In particular, a set of metrics that determines solution quality is found and applied to find a promising neighborhood for solution improvement. Section 4.6 summarizes the findings from the experiment results, and Section 4.7 concludes the chapter with the contributions and recommendations for future work.

4.1 Introduction

Two mathematical formulations have been proposed for the sectorization problem with multi-period demand patterns and time-varying controller staffing (MPVC). Even though both formulations can solve an adequate-size problem, i.e. a 56-node network, to a satisfactory optimality gap within a reasonable timeframe, their performances are not guaranteed for a real world application, e.g. over 1,000 nodes. A major difficulty in solving an integer program such as P_{NFB} comes from the binary variable associated with each edge that is used in the Big-M constraint to indicate whether the edge carries positive flow. The optimal value of the linear program

relaxation provides a very weak lower bound (Sridhar and Park 2000, Magnanti 1995).

To address the computational issue encountered for a realistic-size MPVC, it is desirable to develop a heuristic that can generate high quality solution with time constraints. In the following sections, the heuristic based on large-scale neighborhood search is discussed, and a parallel computing framework to effectively improve a given solution of MPVC is proposed. Several neighborhood selection schemes will be proposed and evaluated. Their performance will then be compared and discussed.

4.2 Heuristic Design

4.2.1 Large-Scale Neighborhood Search

Heuristic approaches are usually proposed for practical purposes to provide solutions in a timely manner. However, the performance of a heuristic depends heavily on neighborhood search. Even in the field of metaheuristics, well-performing methods still rely on defining and changing neighborhoods in order to avoid getting trapped in a local optimum.

For an NP-Hard problem such as MPVC, conventional mathematical programming (MP) seems to be a less desirable choice in practice for producing quick solutions. In recent years, commercial mixed-integer-programming (MIP) solvers have become increasingly efficient. Their customized codes have been effectively used in a heuristic context. A general idea is that if a problem can be decomposed into manageable subproblems, it can then be solved by iteratively defining and solving a subproblem through exact methods coded in the solvers. For example, a subset of

routes in a feasible solution of a vehicle routing problem (VRP) can be re-optimized by solving the associated mathematical program.

The MP-based heuristic also refers to large-scale neighborhood search, in contrast to typical neighborhood search that executes a relatively simple procedure within a relatively small neighborhood e.g. two-swap or three-swap in the travelling salesman problem (TSP). Each application of MP to a subproblem involves solving an optimization problem, so its solution space is generally larger than for other neighborhood search methods.

4.2.2 Neighborhood Selection for MPVC

The neighborhood selection scheme (also solution decomposition scheme) is problem-specific and plays an important role in the effectiveness of the heuristic. A wide range of schemes can be developed to define neighborhoods. Mathematically, Ball (2010) describes such decomposition methods as "row partition of a solution" and outlines a conceptual algorithm presented in Fig 4-1:

Row Partitioning
Input: feasible solution

1. cut out a portion of the solution of manageable size, i.e. a row subset;
2. apply an exact method over the portion cut out in (1);
3. paste the solution obtained in (2) back into the original solution;
4. if new, improved, feasible solution is created then repeat.

Figure 4-1 Row Partitioning Algorithm Illustrated in Ball (2010)

The decomposition scheme relies on solution properties and problem natures. Bent and van Hentenryck (2007) proposed a randomized adaptive spatial decoupling scheme for VRP with time window that iteratively selects and re-optimizes a subset

of vehicle routes. Their scheme is claimed to be adaptive because it depends on the current solution that evolves with any improvement found. To tackle an on-demand air transportation problem based on a multi-commodity flow model, Espinoza et al. (2008) also proposed an optimization-based local search to obtain quality solutions for large-size real-world instances. Any existing solution can be decomposed into jet itineraries, a subset of which defines a neighborhood. They also developed several metrics of solution quality and tested the performance of selection schemes based on metrics and on random choices.

Judging solution quality before selecting neighborhoods has also been applied in Sniezek and Bodin (2006) on a capacitated arc routing problem for residential sanitation collection vehicles. A measure of goodness for a feasible solution was developed that included route design objectives not addressed in the original objective function, so the modeler could address the tradeoffs among various design concerns during local search.

In sum, solution decomposition schemes help identify a partial, manageable solution of poor quality that has the potential for improvement after re-optimization.

For MPVC, a feasible solution consists of the sector boundaries and time-varying controller staffing, so an intuitive decomposition scheme for applying large-scale neighborhood search is to select a group of geographically adjacent sectors, and decouple them based upon sector spatial relations.

The idea is illustrated in Fig 4-2. Suppose we are given a feasible, global solution to MPVC, i.e. sector boundaries with time-varying staffing plans. A group of sectors, representing a local solution, is then selected and constructed as a MIP for re-

optimization. If an improvement after re-optimization is found, then the global solution will be updated accordingly; otherwise, it stays unchanged. These steps are iteratively repeated until the stopping criteria are met.

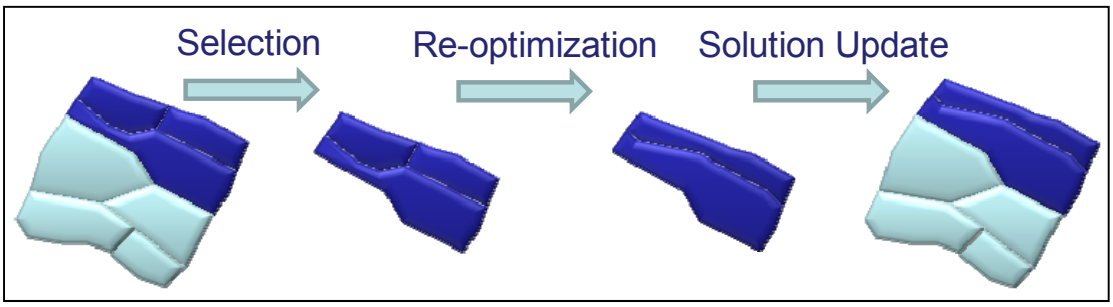


Figure 4-2 Solution Decomposition and Local Improvement

This decomposition approach can be classified as solution improvement heuristic that uses MIP solvers for local optimization, iteratively improving any feasible solution. With this local search approach with spatial decomposition scheme, there are several considerations to be further addressed:

- The size of each sector group (or solution neighborhood) is not arbitrary.
- The neighborhood selection scheme would influence the effectiveness of the heuristic.
- The computation framework is parallelizable to save computation time.

The neighborhood size should be judiciously determined, and it relates to the possibility of being trapped within a local optimum. A larger neighborhood tends to bring significant improvement. However, the neighborhood size also affects how fast a subproblem converges to an acceptable gap in a MIP solver. Although the performance is not guaranteed, it is highly possible that smaller neighborhoods converge quicker than larger ones and thus result in more runs within a given time

limit. There exists a tradeoff between the neighborhood size and satisfactory convergence rate.

In addition, the selection strategy of neighborhoods (or subproblems) determines the efficiency of this heuristic. An intuitive strategy is randomized selection, which randomly picks adjacent sectors for re-optimization. This is intuitive and simple but not necessarily most efficient. Since computation time is precious and should be used efficiently, we will investigate the metrics that measure the potential of improvements from candidate neighborhoods. The metrics being developed will be then applied to a guided selection scheme, and their performance will be analyzed.

The neighborhood selection can be done so that several neighborhoods are selected simultaneously and mutually exclusively. The subproblems formulated for these neighborhoods are separable and independent of one another. The computation framework proposed for this heuristic is thus designed to solve several subproblems in parallel.

4.2.3 Framework of Synchronous Parallel Computation

The parallel computation framework proposed here takes the advantage of currently popular multi-processor, multi-core computing environment, especially since computers nowadays are commonly equipped with one or more multi-core processors. As long as the shared memory is sufficient, each computing core can handle one single execution without much interfering with others. The proposed framework in Fig 4-3 is a synchronous one and requires three main components:

- The Master Module coordinates with all the modules. Its main function is to assign tasks to Slave Modules, retrieve and process local solutions from Slave Modules, and maintain the global solution.
- The Initial Solution Module provides the initial solution for improvement, either from solving the global version of the problem through MIP solvers or from any solution construction heuristic.
- The Slave Module waits for the Master Module to call. Once receiving sector selection information, it constructs and then solves a MIP. When the stopping criteria are met, the Slave Module sends an event message back to the Master Module if the solution is better than the existing one so that the Master Module can maintain the currently best solution.

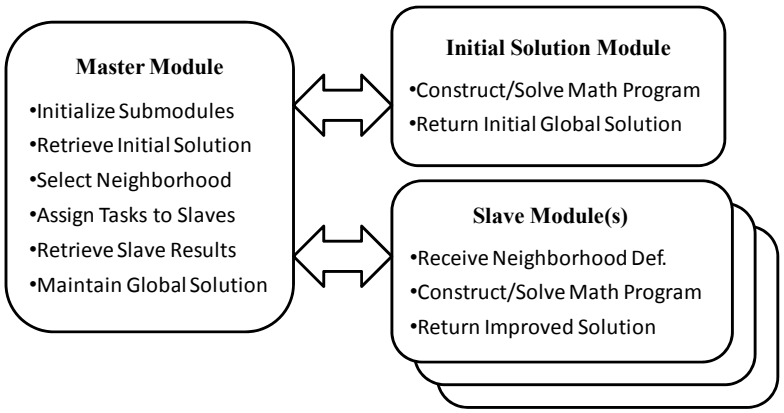
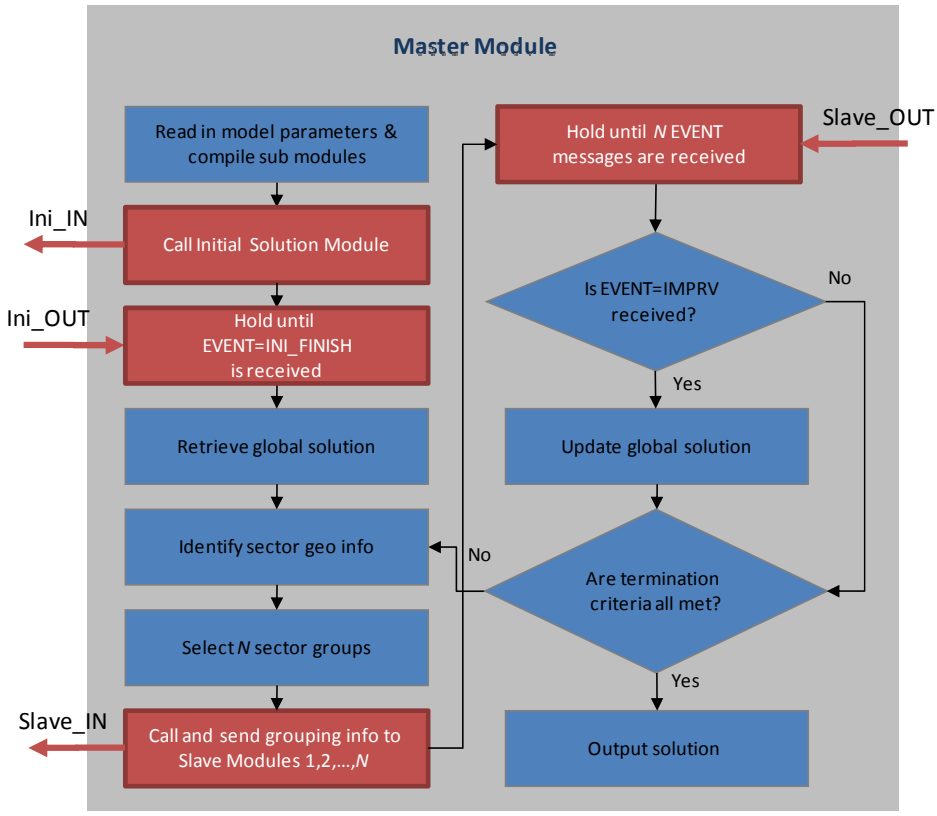


Figure 4-3 Parallel Computation Framework of the MP-based Heuristic for MPVC

The heuristic algorithm is coded under a parallel computation framework in the environment of Xpress-Optimizer whose optimization function is thus integrated. A detailed framework design is described in Fig. 4-4 and customized for the formulation P_{NFB}. Since P_{NFB} requires additional efforts for processing the information of node-sector and link-sector assignments in order to identify sector geographical

contingency, the "**Identify sector geo info**" block in the Master Module is coded accordingly.



Master Module

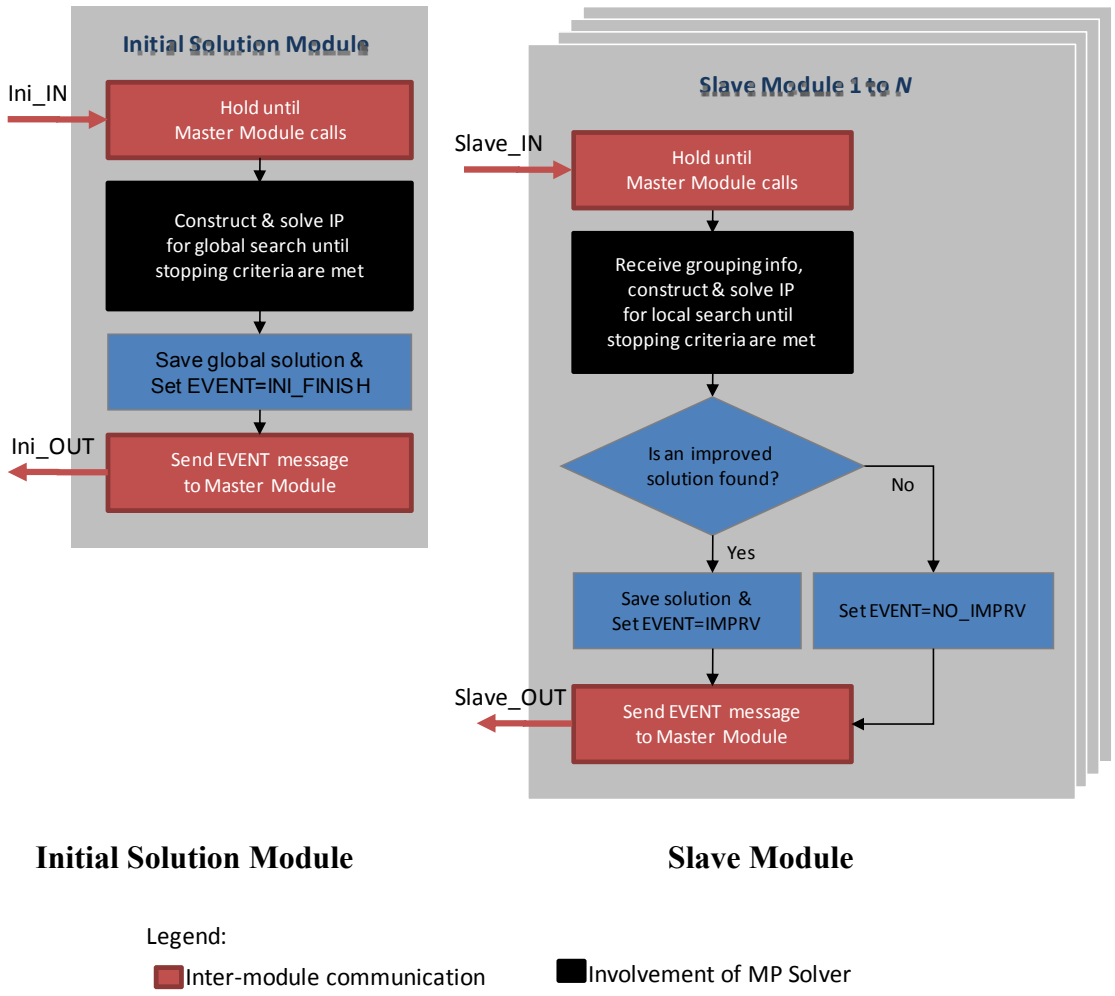


Figure 4-4 Modules of the Proposed Parallel Computation Framework in Xpress

4.3 Random Selection Scheme

In the rest of the chapter, the effectiveness of the proposed heuristic is demonstrated by using different neighborhood selection schemes.

The same dataset as in Chapter 3 is used for the following experiments, which is one day traffic (April 21, 2005) at the Washington DC Center (ZDC). Again, it is assumed that there are 4 periods, each of which has 4 hours, and at most 2 positions can serve sector traffic. The airspace under study is represented by a network of 1,043 nodes with 4,400 Links.

To obtain a comparison basis, we first solve the instance by using Xpress on a Dell PowerEdge 1900 with Intel Xeon 2.66Ghz processor and 12 GB memory and obtain the computation results of different stopping criteria. The solver has been tuned with the best found strategies of branch-and-bound search. The instance is solved and the solution is retrieved after 1 hours, 3 hours, and 24 hours. The results are summarized in Table 4-1.

After solving for one day, the MIP gap is 7.32%. There is no significant improvement found from 3 to 24 hours for both the LP bound and the objective function value. A slight improvement comes from the reduction of "flow cost", which is related to flow alignment with sector shape.

Table 4-1 Global Search Results

Global Search Time		1hr	3 hrs	24 hrs
MIP Results	Objective Value *	2,589,998	2,308,840	2,306,466
	- Controller Cost	900	756	756
	- Flow Cost	969,998	948,040	945,666
	MIP Gap (%) **	17.47%	7.42%	7.32%
Sector Characteristics	No. of Sectors	18	18	18
	No. of Controllers	92	76	76
	No. of 2nd Positions	20	4	4

* Tradeoff coefficient μ=1800
** Use best found bound at 24 hrs (2137550).

For the experiment settings, we use the 1-hour solution as the initial solution for improvement and run the improvements for 3 hours. The first selection scheme is to randomly select the neighborhood, i.e. the sector groups. Because multiple sector groups can be evaluated in parallel, the procedures of the random selection scheme stated below will be repeated until the ideal number of neighborhoods is found.

Random Selection Scheme

Input: A set of candidate sectors L, their adjacency relation $adj(L)$.

Output: A subset of N contiguous sectors S.

Procedures:

1. Set S empty. Select a sector i randomly, where i is in L. Set $L := L \setminus \{i\}$ and $S := S + \{i\}$.
2. If $adj(S) \cap L$ is not empty, then select a sector j randomly, where j is in $adj(S) \cap L$. Set $L := L \setminus \{j\}$ and $S := S + \{j\}$; otherwise, reset L and go back to Step 1.
3. Repeat Step 2 until N sectors have been selected.

There are 4 types of neighborhood size to be considered, i.e. pair, triple, quadruple, and quintuple. The size of neighborhood is an important factor that not only impacts the performance of the proposed heuristic but also determines a suitable number of slave modules to be used in the parallel computation. For example, for the case of a 5-sector neighborhood, the diversity of defining 3 neighborhoods of 5 in a solution of 17 sectors is very limited, and the heuristic could repeatedly evaluate the same set of sector selections without finding any improvement. Thus, only 2 slave modules are called for the quintuple case, and 3 for the rest of the cases.

In addition, the time limit for solving a subproblem also depends on neighborhood size. Larger size would avoid entrapment in a local optimum whereas smaller size could be solvable to reach a satisfactory convergence level. Due to lack of selection diversity, the neighborhood size of 6 or above is not considered since there are few combinations of 6 or more geographically connected sectors.

Table 4-2 summarizes the results for the random selection scheme. The time limit of each iteration is set to vary with neighborhood size – a larger neighborhood implies a larger subproblem and generally requires more time to reach a satisfactory gap.

Since randomness is involved in the neighborhood selection, the experiment is run 10 times on each neighborhood size and the performance metrics are averaged. The overall improvement on the initial solution is around 11%. The "quadruple" size performs better within time constraints, i.e. higher average improvement rate and lower standard deviation, than other neighborhood sizes.

By calculating the number of successful iterations divided by that of total iterations, the success rate of improvement attempts can be evaluated. Smaller neighborhoods have higher success rates.

Larger neighborhoods require more search time, so the number of iterations within the time limit is lower. However, larger neighborhoods define a larger solution space for local search. It is observed that although the "quintuple" size has the lowest success rate, on average its improvement rate per success is the highest among other neighborhood sizes.

Except for "pair", all the neighborhood sizes reach better (lower) MIP gaps than the best found gap (the 24-hour case). In this particular instance and random selection scheme, the "quadruple" size seems to be adequate for obtaining a better solution than the best found exclusively by using Xpress.

Table 4-2 Local Search Results – Random Selection Scheme

Category		Pair	Triplet	Quadruple	Quintuple
Experiment Settings	Global Search Time (hr)	1	1	1	1
	Local Search Time (hr)	3	3	3	3
	No. of Slave Modules	3	3	3	2
	Time Limit per It. (sec)	150	200	300	300
Random Selection Scheme (Case 1plus3)	No. of Iterations	60.50	41.75	29.88	31.50
	No. of Success	52.25	34.00	24.25	14.00
	Succ. Rate (%)	86.31%	81.47%	81.24%	46.48%
	Avg. Impr. per Succ. (%)	0.20%	0.33%	0.49%	0.79%
	Avg. Impr. on Obj Fcn Val (%)	10.36%	11.21%	11.78%	11.01%
	Std. Dev of Impr. (%)	0.80%	0.87%	0.71%	0.80%
	MIP Gap (%)	7.92%	7.04%	6.45%	7.25%

4.4 Defining Metrics for Measuring Solution Quality

As shown in the previous section, the proposed heuristic with a random selection scheme found better MPVC solutions within 4 hours than the best found exclusively by the solver within 24 hours. However, computation time is expensive – the random selection might still not be efficient in the sense that the computation effort would be spent on the subproblems that are very close to the optimality while those potentially improvable subproblems were seldom visited. In the literature, neighborhood selection based on metrics developed for measuring solution quality has demonstrated its potential in providing proper guidance of finding a neighborhood that can be improved. To apply the proposed heuristic more effectively and efficiently, there is a need to develop the metrics to identify the quality of solutions.

One might think that the original objective function itself can be applied to judge the quality of a decomposed solution; however, it is not sensitive enough to identify a promising neighborhood, as indicated in Sniezek and Bodin (2006). Therefore, we propose two metrics for evaluating the solution quality for a sector, described as follows:

- **Metric 1 – Capacity Surplus**: The difference between the capacity provided by controller staffing and the demand actually served.
- **Metric 2 – Deviation of Link Cost to MST Value**: The difference between the cost of selected links and the cost of the minimum link-cost spanning tree.

The first metric represents the potential to better utilize the controller resources by expanding or reducing sector size and is calculated as follows:

- Metric 1 for sector $i = \sum_{t=1}^{T}\sum_{p=1}^{P}\left(U_p \hat{g}_{i,p}^{t} - \hat{s}_i^t\right)$

  where $\hat{g}$ and $\hat{s}$ are the current solutions of controller staffing and sector demand, respectively of a candidate sector i, which might be selected for improvement.

It is proved in Lemma 3-1 that at optimality the links that form a sector constitute a minimum spanning tree. The second metric based on this solution property helps identify whether the optimality condition for a subproblem is achieved.

- Metric 2 for sector $i = \sum_{\substack{(j,k)\in E \\ j,k \text{ in sector } i}} c_{jk}\hat{w}_{jk} - MST_i$

  where $\hat{w}$ is the current solution of link variables in sector i and MST_i is the value of minimum link cost spanning tree of all the nodes assigned to sector i.

To validate the effectiveness of the proposed solution metrics, we investigate the statistical relation between heuristic performance metrics and solution quality metrics. The sector boundaries in the 1-hour solution illustrated in Fig. 4-5 are used as study object. The "pair" neighborhood size is considered, and all possible pairing choices are enumerated. The sector boundaries are improved by using the proposed heuristic with respective sector pairs for one iteration. Each of the subproblems associated with sector pairs is solved to the optimality (since the size of each subproblem is relatively small), and their performance metrics are summarized in Table 4-3.

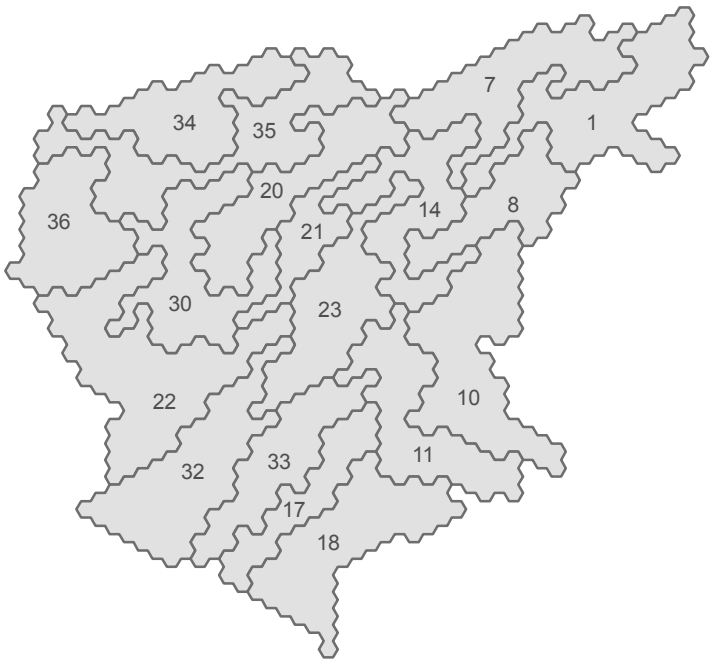


Figure 4-5 Sector Boundaries from the 1-Hour Solution

Table 4-3 Relation between Local Search Performance and Solution Quality Metrics

Pair Index	Sector		Performance		Metric 1	Metric 2
	1st	2nd	Obj. Fcn. Val. Improv. Rate	Flow Cost Improv. Rate	Capacity Surplus per Sector	Dev. of Link Cost to MST Value
1	1	7	26.67%	5.22%	4455.99	5731.30
2	7	14	21.88%	5.06%	3235.99	5151.30
3	1	8	20.84%	2.50%	4139.99	1685.00
4	7	20	17.21%	4.32%	3970.00	4238.30
5	8	14	15.23%	2.09%	2919.99	1105.00
6	34	35	13.49%	1.56%	3532.01	1084.70
7	20	21	12.80%	0.28%	3659.01	386.70
8	21	23	12.30%	2.76%	3585.00	2755.10
9	23	33	11.76%	1.90%	4203.50	2671.70
10	8	10	11.13%	1.18%	3906.49	1299.70
11	21	30	9.31%	5.23%	2387.50	3188.30
12	14	21	8.42%	2.16%	2925.01	1300.00
13	1	14	7.25%	2.32%	2807.00	2713.00
14	35	36	6.80%	0.76%	2267.00	643.70
15	10	14	5.72%	0.86%	2573.51	2328.70
16	21	22	5.67%	1.11%	2735.00	985.10
17	20	35	5.66%	1.51%	2711.01	683.00
18	23	32	3.70%	1.79%	2661.00	2909.70
19	17	33	3.23%	-0.95%	3290.00	150.00
20	11	33	3.05%	-1.00%	3067.49	149.90
21	32	33	2.84%	-0.33%	3957.50	538.30
22	14	23	1.61%	3.98%	2247.00	3588.70
23	20	30	1.38%	4.43%	1783.50	3108.30
24	30	36	1.30%	3.28%	1339.49	3068.30
25	30	35	1.29%	3.11%	1439.50	3485.30
26	22	30	1.16%	3.04%	859.48	3707.00
27	11	23	1.12%	3.08%	1770.99	2521.70
28	22	23	1.01%	2.66%	2056.99	3273.40
29	11	14	0.48%	1.36%	1110.99	1066.60
30	14	20	0.43%	1.42%	2321.01	1220.00
31	10	11	0.43%	1.19%	2097.49	1262.00
32	22	32	0.39%	0.98%	1811.00	1139.70
33	22	36	0.34%	0.94%	1686.99	865.40
34	17	18	0.03%	0.08%	1070.50	40.10
35	11	18	0.02%	0.05%	848.00	40.00
36	11	17	0.00%	0.00%	857.49	149.90

Fig. 4-6 shows the relation between Metric 1 and the improvement rate of the objective function value. The estimated function shows a nonlinear, increasing trend, and its r-squared value suggests the statistical significance.

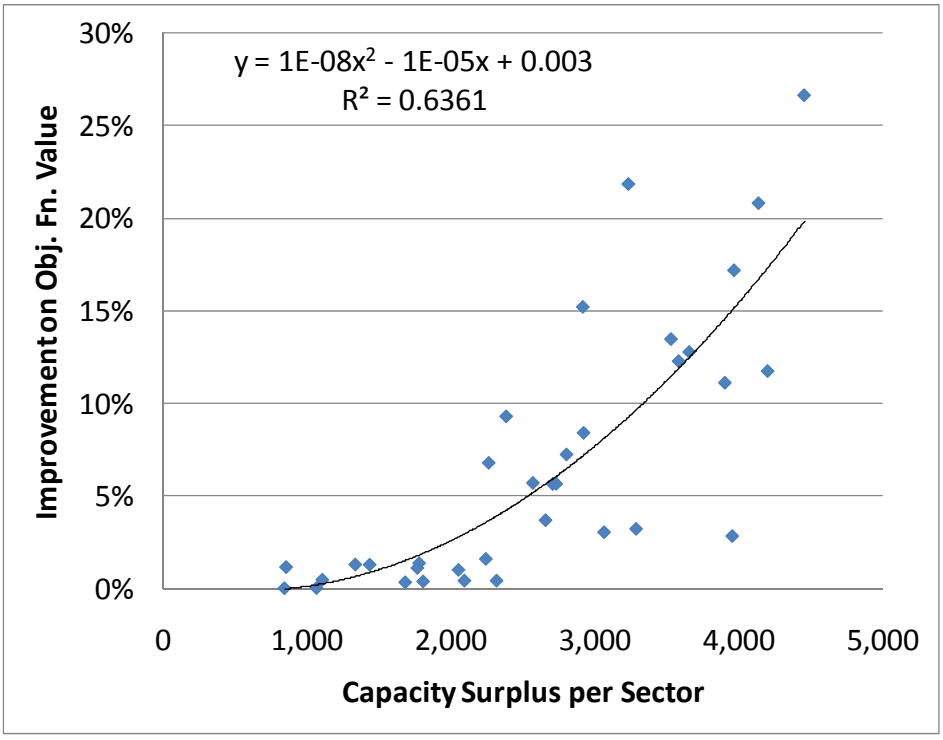


Figure 4-6 Metric 1 vs. Local Search Performance

Fig. 4-7 illustrates the relation between Metric 2 and the improvement rate on the link cost (flow alignment) part of the objective function value. Its statistical significance is even more obvious than that of Metric 1. While MPVC optimizes a multi-objective function, there exists a tradeoff between the controller cost and the flow alignment penalty cost. Since the controller cost is the dominant objective, we occasionally observe a situation in which controller cost is reduced by sacrificing the flight alignment objective. That explains the data points below the horizontal axis in Fig. 4-7.

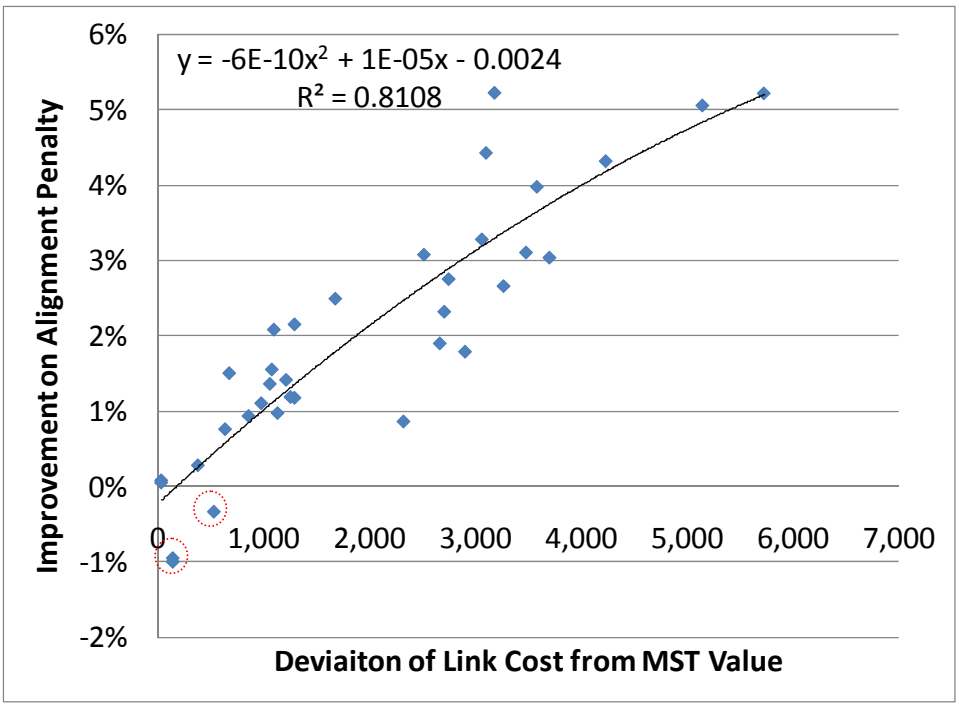


Figure 4-7 Metric 2 vs. Local Search Performance

4.5 Metric-Based Selection Schemes

4.5.1 Scheme Development

In this section, two selection schemes based on the proposed solution quality metrics are proposed, which mainly differ in selecting the first sector in the sector group. The first metric-based selection scheme is described as follows and will be repeated until the ideal number of neighborhoods is found:

Metric-based Selection Scheme (I):

Input: A set of candidate sectors L, their adjacency relation $adj(L)$, A list of metric value v_i associated with each candidate sector i.

Output: A subset of N contiguous sectors S.

Procedures:

1. Set S empty. Select a sector i randomly, where i is in L. Set $L := L \backslash \{i\}$ and $S := S + \{i\}$.
2. If $adj(S) \cap L$ is not empty, then select a sector $j = \arg\max_k \{v_k \mid k \in adj(S) \cap L\}$. Set $L := L \backslash \{j\}$ and $S := S + \{j\}$; otherwise, reset L and go back to Step 1.
3. Repeat Step 2 until N sectors have been selected.

To cope with two solution quality metrics, we use Metric 1 for the first half of the search span and Metric 2 for the rest. This is done because we observe that larger improvements always result from reducing controller shifts (or efficiently utilizing provided sector capacity). Since Metric 1 identifies where the provided capacity might not be well utilized, it is an effective strategy for exploring the neighborhoods that have the potential to increase capacity utilization. After the heuristic runs for a while, Metric 2 will help identify the neighborhoods where the flow alignment objective can be improved.

Another way adopted to increase selection diversity is using a Taboo list that records the first sectors for all the slave modules in the previous iteration so that the neighborhood selection does not explore the previously visited sector groups.

One might argue why sectors are not selected sectors purely by metrics. Since the proposed metrics can by no means guarantee solution improvement, it is observed during the previous experiment performed that such strategy will result in repeatedly

evaluating a small range set of neighborhoods and being trapped into a local optimum. The randomness imposed on choosing the first sector is intended to increase the diversity of starting a neighborhood selection, thus increasing the possibility of finding good solutions.

We modify the selection method of the first sector and propose another scheme that incorporates solution quality metrics into finding the first sector by using a weighted random number. The weight is the value of the selected metric normalized between 0 and 1. If a sector has a poorer (higher) metric value, its likelihood of being chosen is higher. Randomness is still kept for the same reason. The second metric-based selection scheme is described as follows and will be applied to select neighborhoods in the same manner:

Metric-based Selection Scheme (II):

Input: A set of candidate sectors L, their adjacency relation *adj(L)*, A list of *normalized* metric value v_i associated with each candidate sector i.

Output: A subset of N contiguous sectors S.

Procedures:

1. Set S empty. Select a sector $i = \arg\max_k \{v_k \times rn_k \mid k \in L\}$, where rn_k is a uniform random number between 0 and 1. Set $L := L \backslash \{i\}$ and $S := S + \{i\}$.
2. If $adj(S) \cap L$ is not empty, then select sector $j = \arg\max_k \{v_k \mid k \in adj(S) \cap L\}$. Set $L := L \backslash \{j\}$ and $S := S + \{j\}$; otherwise, reset L and go back to Step 1.
3. Repeat Step 2 until N sectors have been selected.

The metrics are computed in the block of "**Identify sector geo info**" in the parallel computation framework in Fig. 4-4(a) right after the Master Module receives and

processes re-optimization results. The proposed selection schemes are then incorporated in the block of "**Select N sector groups**".

4.5.2 Performance Comparison of Neighborhood Selection Schemes

Two selection schemes are applied in the heuristic under the parallel computing framework. The experiment settings are the same as the random selection scheme. The 1-hour solution serves as the initial solution to improve upon. The results are summarized in Table 4-4.

Table 4-4 Local Search Results – Metric-based Selection Schemes

	Category	Pair	Triplet	Quadruple	Quintuple
Experiment Settings	Global Search Time (hr)	1	1	1	1
	Local Search Time (hr)	3	3	3	3
	No. of Slave Modules	3	3	3	2
	Time Limit per It. (sec)	150	200	300	300
Metric-Based Selection Scheme I (Case 1plus3)	No. of Iterations	60.50	41.63	30.13	31.88
	No. of Success	45.75	34.88	22.13	12.00
	Succ. Rate (%)	75.49%	83.87%	73.52%	37.73%
	Avg. Impr. per Succ. (%)	0.24%	0.34%	0.53%	0.97%
	Avg. Impr. on Obj Fcn Val (%)	10.88%	11.83%	11.78%	11.69%
	Std. Dev of Impr. (%)	0.54%	0.59%	0.46%	0.65%
	MIP Gap (%)	7.39%	6.39%	6.44%	6.54%
Metric-Based Selection Scheme II (Case 1plus3)	No. of Iterations	58.75	41.875	30.375	32.25
	No. of Success	47.375	33.75	21.5	11.75
	Succ. Rate (%)	80.87%	80.72%	70.87%	36.55%
	Avg. Impr. per Succ. (%)	0.23%	0.36%	0.58%	1.02%
	Avg. Impr. on Obj Fcn Val (%)	11.05%	12.28%	12.39%	11.93%
	Std. Dev of Impr. (%)	0.33%	0.41%	0.36%	0.47%
	MIP Gap (%)	7.21%	5.92%	5.80%	6.29%

Overall, the Scheme II, which uses the weight random number to find the first sector, improves not only the overall improvement rates but also all other performance statistics.

For the MIP gap as well as the average improvement rate on the objective function values, two metric-based selection schemes outperform the random one, as illustrated in Figs. 4-8 and 4-9. Mid-size neighborhoods, i.e. triplets and quadruples, still have better performance than others.

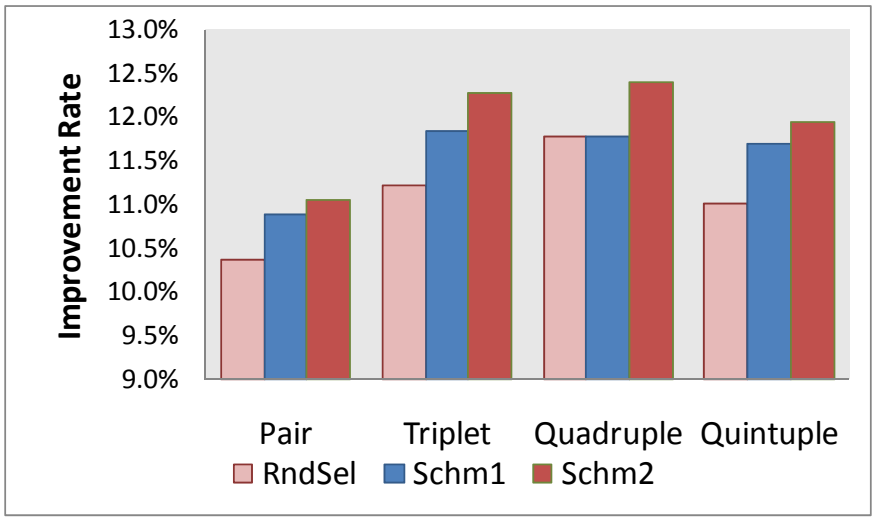


Figure 4-8 Average Improvement Rate of Three Selection Schemes

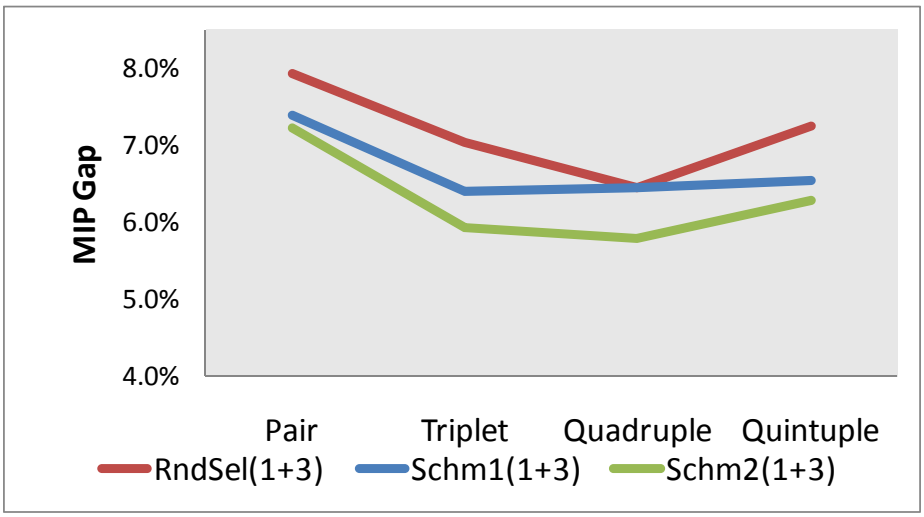


Figure 4-9 Average MIP Gap of Three Selection Schemes

The decreasing standard deviation of the improvement in Fig. 4-10 also suggests that the behavior becomes more and more consistent from purely random to metric-based selections as the computational efforts are concentrated on the promising neighborhoods.

Compared with the results of random selection scheme (in Table 4-2), the success rate decreases, but the improvement per success increases. The increasing step size per improvement illustrated in Fig. 4-11 results from the fact that the metric-based schemes lead the search effectively to promising neighborhoods.

By observing the improvement over the 3-hour search span, Scheme II approaches the final solution in an earlier stage. In Figs. 4-12 and 4-13, the improvement of the best run for each neighborhood size is visualized against time horizon for both metric-based schemes, respectively. For Scheme I, there is still significant improvement for all the neighborhood sizes after running the heuristic for 1 hour. For Scheme II, all neighborhood sizes except for the "pair" approach to final solutions in the early stage of the timeframe. This quick convergence behavior of Scheme II suggests that in order to reach a satisfactory gap within stringent time limitations, Scheme II would be a promising choice.

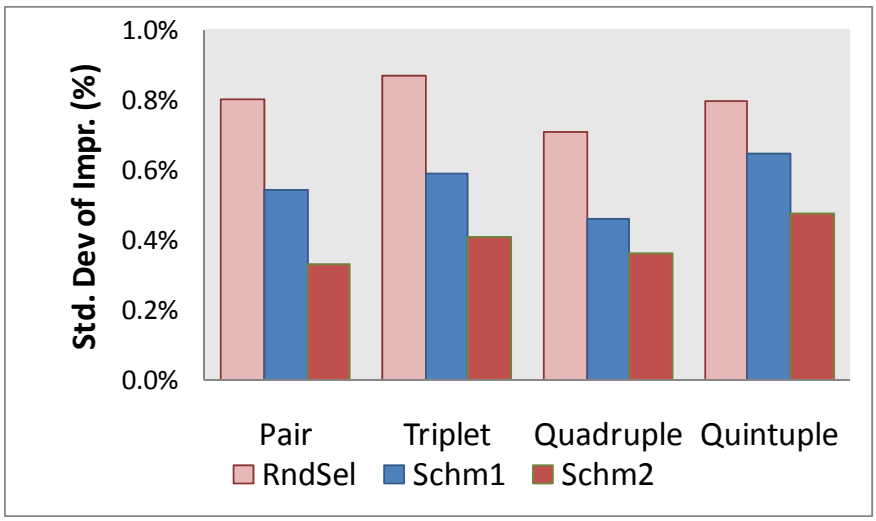


Figure 4-10 Standard Deviation of Improvement Rates

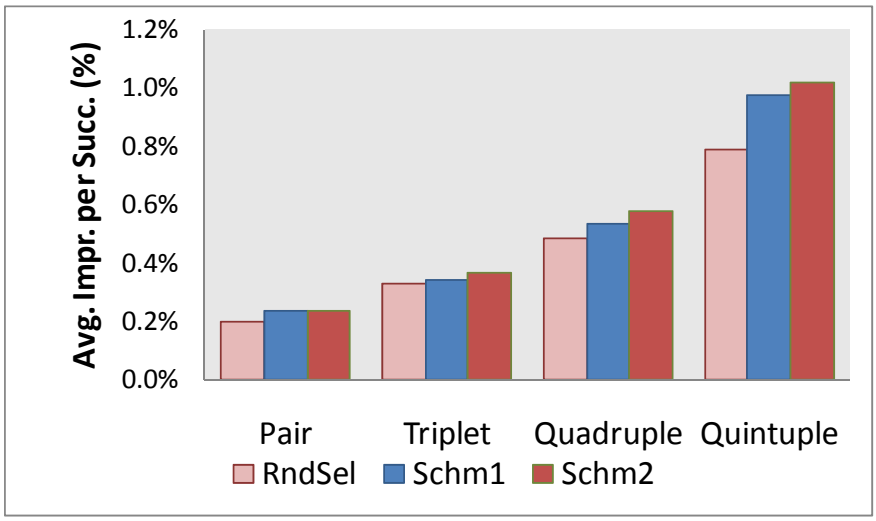


Figure 4-11 Average Improvement Rater per Successful Iteration

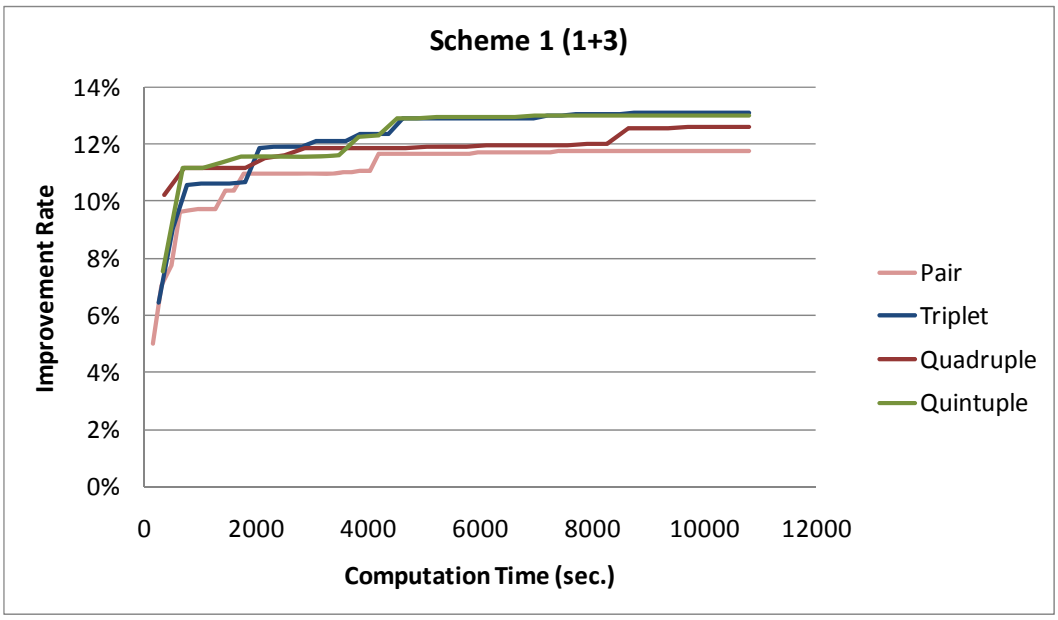


Figure 4-12 Improvement Rate over Time of Best Improvement – Scheme I

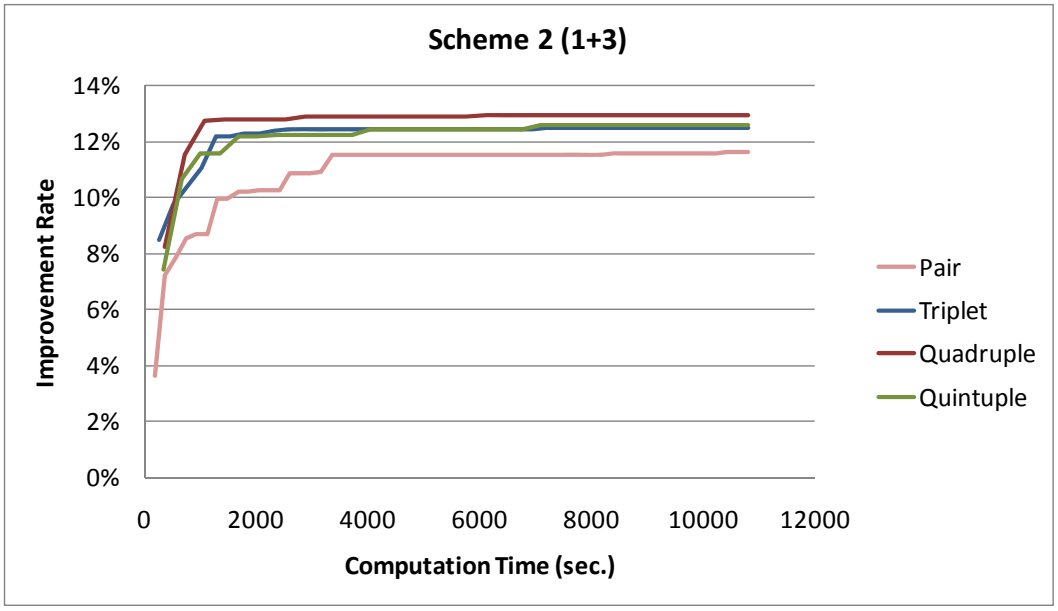


Figure 4-13 Improvement Rate over Time of Best Improvement – Scheme II

4.5.3 Sensitivity to the Quality of Initial Solution

Obtaining a good initial solution through solving a global problem with an exact method is time consuming. Especially for a large-scale problem such as a realistic MPVC instance, there is no guarantee that the computational time and provable

solution quality are proportional. Yet it is unclear that a good initial solution will help the improvement heuristic find a good solution. Further study is desirable to determine whether spending time on getting a good initial solution or on improving an arbitrary feasible one.

We run the proposed heuristic on the 3-hour solution, i.e. the solution obtained after 3 hour global search, which is much better than the 1-hour solution used in previous experiments. With the same settings and using Scheme II, the experiment results are summarized in Table 4-5. The average improvement on the objective function value is very limited, from 0.33% to 1.8%.

A good initial solution might not be suitable for a smaller neighborhood because it easily leads the search into a local optimum. In Fig. 4-14, the success rate is generally lower than that of a worse initial solution, except for the quintuple case. Re-optimization over a large neighborhood would help the search escape from such a local optimum.

In Fig. 4-15, the performance on the MIP gap is again compared with the results of a worse initial solution. The best average performance of all the experiments occurs in the quintuple case with a good initial solution. In addition, the downward trend also suggests that with a good initial solution, the choice of neighborhood definition favors large neighborhoods.

Table 4-5 Local Search Results – Sensitivity of a Good Initial Solution

Category		Pair	Triplet	Quadruple	Quintuple
Experiment Settings	Global Search Time (hr)	3	3	3	3
	Local Search Time (hr)	3	3	3	3
	No. of Slave Modules	3	3	3	2
	Time Limit per It. (sec)	150	200	300	300
Metric-Based Selection Scheme II (Case 3plus3)	No. of Iterations	67.90	43.14	30.71	32.10
	No. of Success	28.30	16.71	11.29	19.10
	Succ. Rate (%)	41.68%	38.74%	36.74%	59.50%
	Avg. Impr. per Succ. (%)	0.01%	0.05%	0.13%	0.09%
	Avg. Impr. on Obj Fcn Val (%)	0.33%	0.88%	1.47%	1.80%
	Std. Dev of Impr. (%)	0.09%	0.50%	0.47%	0.59%
	MIP Gap (%)	7.11%	6.60%	6.04%	5.72%

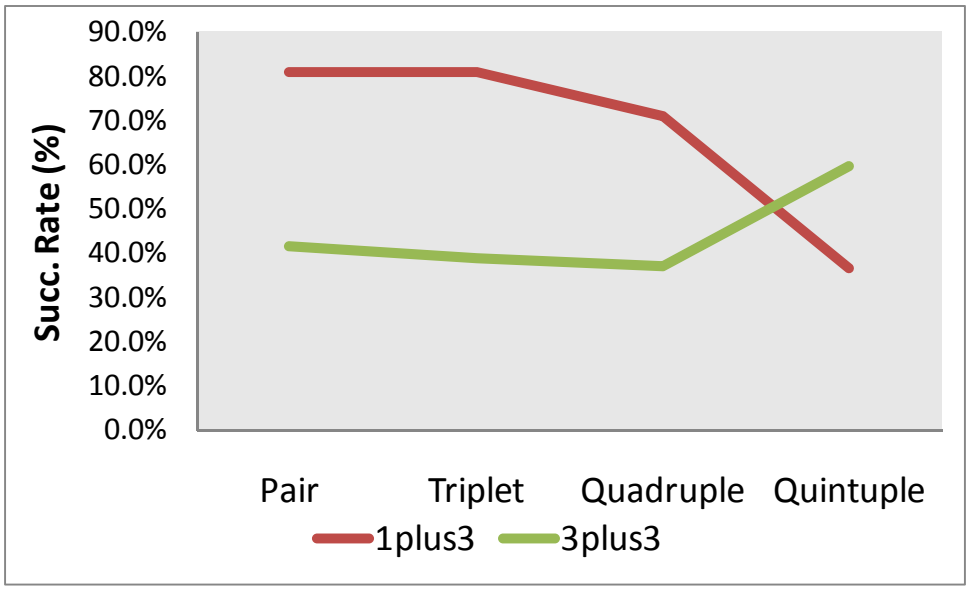


Figure 4-14 Success Rate for Two Initial Solutions

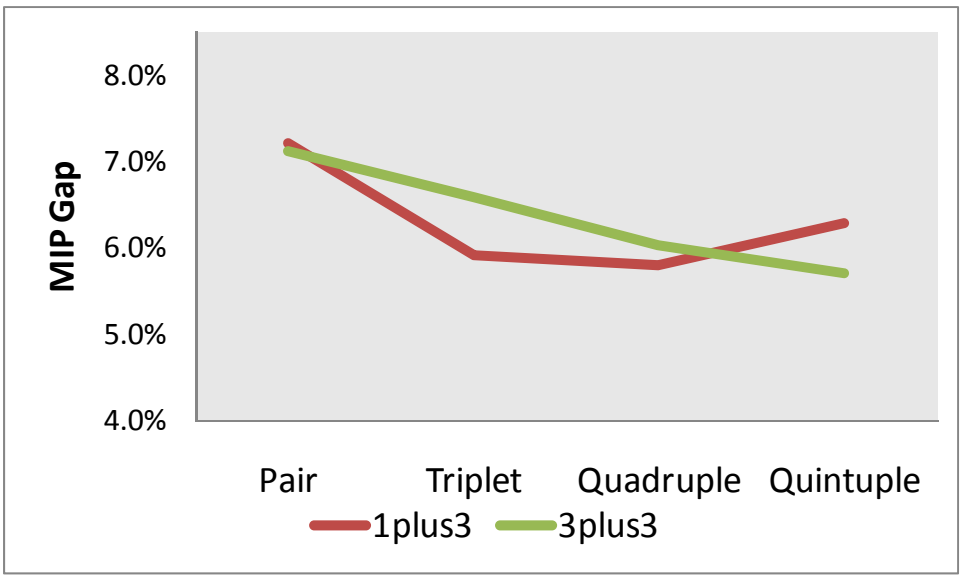


Figure 4-15 MIP Gap for Two Initial Solutions

4.6 Summary and Contributions

The proposed heuristic has demonstrated its capability of improving solutions. Various factors that could affect the heuristic's performance, such as the size and selection of neighborhoods, and initial solution quality, have been analyzed and their impacts on the heuristic's performance have been examined through computational experiments.

In Table 4-6, the best sectorization (solution) found with each experiment is presented. Metric-based selection schemes still perform better. While the commercial solver cannot improve the MIP gap of 7.32% within 24 hours (the result of Global24), Scheme II initialized with the 3-hour solution reaches the lowest (best) gap of 5.11%. The difference of 2.21% might seem fairly small; however, if the solutions are further interpreted with physical meanings, a poor solution, i.e. "Global24", not only uses controller resource inefficiently but also produces inferior sector boundaries design.

Among the heuristic solutions, three of four employ 17 sectors and 68 controller positions over the planning horizon. This implies that the only differences among these three solutions are the flow alignment penalty and thus sector shapes.

Depending on the initial solution quality, the quadruple and quintuple neighborhoods are promising choices for solution improvement. The result echoes the finding in Espinoza et al. (2008) that choosing neighborhoods strategically based on quality metrics favors larger neighborhoods.

Table 4-6 Best Solutions Found by Individual Experiments

	Experiment	Neighborhood Size	MIP Gap	No. of Sectors	No. of Total Controllers	No. of 2nd Controllers
Solver Result	Global24	-	7.32%	18	76	4
Heuristic Results 1-hr Initial Solution	Rnd Schm	quadruple	5.62%	16	64	10
	Scheme I	quintuple	5.34%	17	68	5
	Scheme II	quadruple	5.19%	17	68	5
3-hr Initial Solution	Scheme II	quintuple	5.11%	17	68	5

The main contributions of this chapter can be summarized as follows:

- Developed a large-scale neighborhood search heuristic for MPVC that can find a significant improvement upon the best solutions found through the MIP solver's global search.
- Designed for the proposed heuristic a parallel computation framework to increase evaluations within time constraints and potentially expedite the solution evolution progress.

- Developed solution quality metrics that can help identify the neighborhoods for re-optimization and validate their effectiveness through numerical analysis.
- Examined the impact of neighborhood sizes on heuristic performance and proposed three neighborhood selection schemes, two of which incorporate solution quality metrics and demonstrate their effectiveness on the proposed heuristic. Metric-based selection schemes have higher performance consistency as a result of higher improvement per success and smaller standard deviation.
- Analyzed the sensitivity to the quality of initial solutions; this suggests that a good initial solution would help in finding a good improvement if the neighborhood size is sufficient to allow escape from a local optimum.

Chapter 5: Quantifying the Relation between Traffic and Controller Staffing Decisions

5.1 Introduction

Maintaining a safe, efficient operational environment is the ultimate goal for air traffic control. It requires advanced technology, flexible air traffic management policies and well-trained and experienced air traffic controllers. When safety is ensured, efficiency can be improved by consuming fewer resources or increasing utilization of inputs. We have considered in Chapter 3 how enroute airspace can be sectorized while jointly considering time-varying demand patterns and controller staffing decisions. One of the building blocks of our sectorization models is the assumption that without compromising safety, controller staffing can efficiently vary with workload. The quantification of such a relation is desirable, although the decision about adding another controller position is indeed difficult to simulate since it varies with sector characteristics, facility cultures, and individual differences among controllers. In addition, radar hit count was the only metric used previously to represent sector traffic conditions and it was assumed to be additive. When each time period is long enough, e.g. 2 hours, it would be appropriate to assume linear additivity of the hit count and its usefulness as a good approximation of traffic complexity. However, the challenge of applying these assumptions emerges when the time interval of interest changes from hours to minutes.

Another challenge is that such a relation might not be easily observed from the empirical data. The difficulty might arise if any of the following situations exists:

- Staffing redundancy is planned due to safety concerns – staffing efficiency might not be the main goal of air navigation service providers.
- The traffic metrics used in the staffing standards, if followed, under/over-estimate the true workload experienced by air traffic controllers.
- Staffing decisions do not closely respond to traffic due to long look-ahead or cut-off time for adding/subtracting control positions.
- Other labor contract issues, such as required minimum working hours, do not take into account traffic condition.

In this chapter, statistical analysis is conducted for quantifying the relation between traffic conditions and controller staffing decisions. Not only is it necessary to objectively verify the assumption of the traffic-staffing relationship but also essential to extend the sectorization models to the Sector Combination Problem, which models sector combination/split activities and will be introduced in the next chapter. An ordinal probit regression will be applied to predict the probability of using multiple controller positions, given sector traffic characteristics. Moreover, additional traffic complexity metrics will be explored beyond the count of radar hits. It is expected that by determining the significance of the estimated models, the statistical relation targeted will be identified and used to decide controller staffing.

This chapter is organized as follows. In Sections 5.2 and 5.3, the proposed methodology and the sources of proper data to make a good estimation will be discussed. In Section 5.4, the ordinal regression will be specified. To address the differences among sectors, each sector will be treated as a subject and estimated individually. In Section 5.5, the estimation results will be interpreted and discussed.

In Section 5.6, several concluding remarks and suggestions for further extensions will be made.

5.2 Proposed Approach

Sectors and their associated controller positions, i.e., radar, radar associate, handoff, and flight data, are designed to handle daily traffic. Since we would like to approximate the decision about adding another controller position, instead of using human-in-the-loop (HITL) analysis and subjective rating approaches to capture controllers' stress level and workload, the use statistical methods is proposed for objectively quantifying the relation between controller positions and sector traffic, at an aggregate level.

As its theoretical properties have been reviewed in Chapter 2, the ordinal probit model is suitable for predicting categorical, ordinal response variables, such as using 1, 2, or more controllers per sector. To approximate the staffing decisions directly from traffic characteristics, the ordinal probit regression is then applied to predicting the ordinal, categorical staffing decision in a specified period, i.e. n controllers per sector, where $n \in \{1,...,N\}$. Assume y^* is the latent variable, which ranges from $-\infty$ to ∞ and is unobservable. What is observed is the staffing decision n, and the measurement equation (5-1) for the ordinal regression model is assumed to be:

$$n = \begin{cases} 1 & \text{if } -\infty \le y^* \le \mu_1 \\ 2 & \text{if } \mu_1 \le y^* \le \mu_2 \\ \dots & \\ N & \text{if } \mu_{N-1} \le y^* \le \infty \end{cases}, \qquad (5\text{-}1)$$

where $\mu_1,...,\mu_{N-1}$ are the threshold values to be estimated for each category. The latent function is then specified as:

$y^* = \mathbf{W}\boldsymbol{\beta} + \varepsilon$,

where $\mathbf{W}$ is the covariate vector, and $\boldsymbol{\beta}$ is the coefficient to be estimated.

If the error term ε belongs to the normal distribution, the probabilities of decision categories are estimated by using the covariate values in the measurement equation and taking the inverse of the normal distribution function, e.g. Prob($n \le 2$) = $\Phi(\mu_2 - \mathbf{W}\boldsymbol{\beta})$. The probabilities for individual categories P_n can then be derived by taking the differences of the cumulative probabilities for the groups in order, e.g. $\mathrm{P}_1 = \Phi(\mu_1 - \mathbf{W}\boldsymbol{\beta})$, $\mathrm{P}_2 = \Phi(\mu_2 - \mathbf{W}\boldsymbol{\beta}) - \Phi(\mu_1 - \mathbf{W}\boldsymbol{\beta})$, …, $\mathrm{P}_N = 1 - \Phi(\mu_{N-1} - \mathbf{W}\boldsymbol{\beta})$. In other words, the probability for the first category is the first cumulative probability; the probability for the second category is the second cumulative probability minus the first; and so on. Then the prediction can be made by choosing the category with the highest probability.

5.3 Data Collection

Common practices in sector/controller activities related to this study that have been observed in an enroute center are:

- Additional controller position(s) will be added when sector traffic is increasing.
- When a sector is busy, a portion of its control area will be split based on FPAs and designated to a neighboring sector.

- During quiet periods, the control areas of several sectors will be designated to a particular sector managed by one or a team of controller(s).

The above phenomena mainly arise from two decisions, on controller staffing and on sector combination/split. In order to best investigate these decisions with a statistical approach, we obtain the empirical data from three major sources:

- Controller work history (Cru-X/ART) that records the employees' sign-on/-off of control positions of each sector. (FAA, 2008)
- Radar records of aircraft per minute from ETMS (Enhanced Traffic Management System) database (FAA, 2009). Each radar hit record contains flight ID, time stamp, longitude, latitude, altitude, speed, etc.
- Airspace definition from ETMS that describes the boundary points of FPAs.
- Dynamic FPA combination information that records the starting and ending time of an FPA being assigned to other than its home sector.

Due to data availability, three days in July 2007 of ZMP (the Minneapolis center), i.e. July 2-4, 2007, are selected for all the following analyses. For center-wide statistics, several figures are used to illustrate a few important observations among traffic, staffing, and sector combination activities. Fig. 5-1 describes the number of aircraft handled and active sectors per 15 minutes, demonstrating a recurrent peaking pattern. Corresponding to temporal variation of traffic, sectors will be staffed or handed over to adjacent one(s). Here, an "active" sector means it has positive number of controllers serving traffic and is identified through examining controller work history – if all the positions of a sector is closed or designated to another, that sector is not considered an active sector.

For an active sector, there might be more than one control position staffed to serve traffic. Fig.5-2 depicts the temporal changes in the number of sectors and associated positions. A stacked bar shows in a 15-minute interval the number of sectors that used 0, 1, 2, or 3 controllers. As traffic increases during the day, the number of multiple-controller sectors increases.

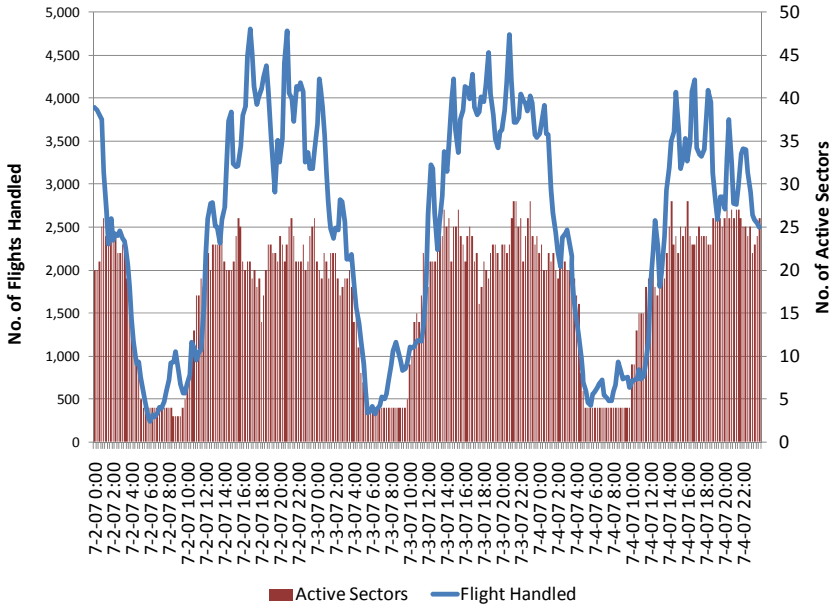


Figure 5-1 Number of Aircraft Handled vs. Active Sectors per 15 Minutes at ZMP during 07/02/2007~07/04/2007

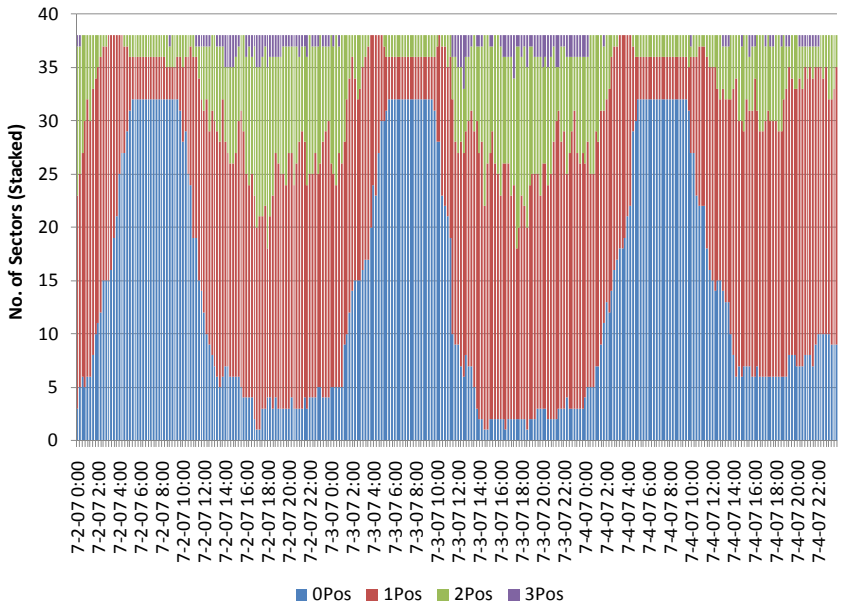


Figure 5-2 Number of Sectors with 0, 1, 2, and 3 Control Positions per 15 Minutes at ZMP During 07/02/2007~07/04/2007

The dynamic FPA combination activities are also of interest. At ZMP, some sectors have up to 9 FPAs while some have only 1. To illustrate an aggregate behavior, the number of FPAs per sector is calculated per 15-minute period. In Fig. 5-3, the blue horizontal line represents a baseline of the average FPAs per sector without any combination. The red line illustrates this ratio with consideration of FPA activities, that is, if at a 15-minute period a sector has 3 "default FPAs" and also receives additional 3 FPAs from its neighbor sectors, it is considered to have 6 FPAs at that period. For each period an average is then taken for all the active sectors. (Note that all the inactive or closed sectors are excluded from the calculation.) It can be observed that during quiet periods the average number increases while during the peak periods, the average number decreases toward the baseline value, which means fewer FPA combination activities. However, this does not mean the FPA combination strictly follows the traffic pattern, because a temporary imbalance of workload during

peak periods might still trigger the designation of one or more FPAs of a busy sector to adjacent sectors, resulting in more combination activities.

The dynamic FPA combination activities define the effective control area served by controllers and should be taken into account in measuring the actual workload experienced by controllers, especially when researchers use various sources to compute sector traffic complexity metrics. Failing to include FPA activities in defining sector boundaries would distort estimates of controller workload.

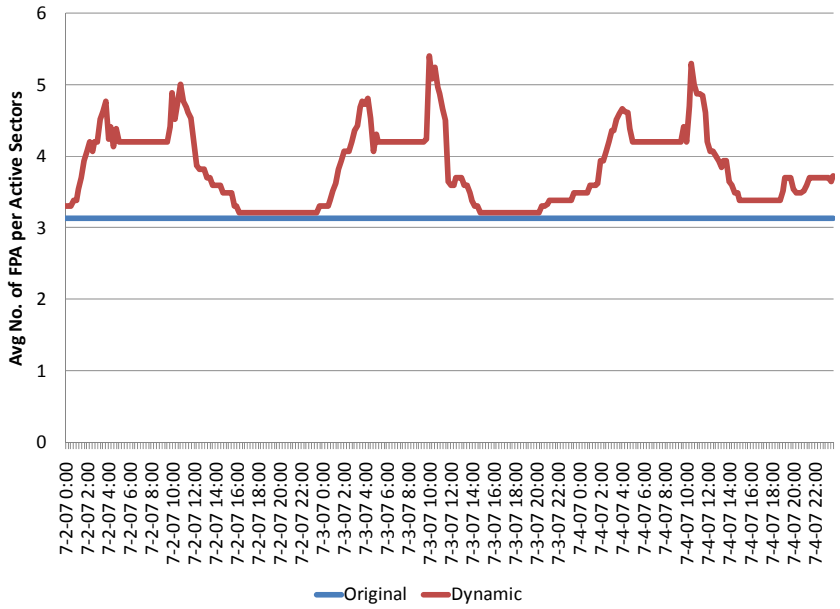


Figure 5-3 Average Number of FPAs per Sector per 15 Minutes at ZMP during 07/02/2007~07/04/2007

5.4 Model Construction and Estimation

5.4.1 Data Processing

The categorical dependent variable considered specifically in this study is the number of open controller positions in a defined period. Cru-X/ART is the official time and attendance system for signing in/out for a shift as well as signing on/off positions

(FAA, 2010). Its data can be further processed to generate sector position counts, i.e. the number of the positions used in each defined bin for each sector.

To quantitatively describe traffic intensity, we consider several variables that correlate to three general categories of controller workload proposed by Delahaye (1995):

- For monitoring workload, the radar hits of each FPA are counted. The hit count can be easily aggregated to sector-wide or center-wide metrics since it is additive. Notably, an alternative traffic metric is the flight count, which is defined as the number of the flights that are handled in an FPA, so by definition the flight count is not additive. We choose the hit count over the flight count because not only does it correlate highly with flight count but also it contains flight dwelling information, that is, the more flights and the longer their stay in a sector, the more hits there are.
- For handover workload, we simply count the number of flights crossing between each pair of FPAs. For aggregation into the sector based metric, if two FPAs are in the same sector, then the crossing count between these two FPAs will not be counted.
- For conflict resolving workload, aircraft conflicts are barely observed from post-operational data since they presumably have been resolved already to avoid collisions. An extensive modeling effort is required to predict conflicts from historical data that involves the extrapolation of flight heading based on previous trajectories, e.g. Hu et al. (1999), Prandini et al. (2000), Sherali et al. (2003). Conflict prediction or detection is beyond the scope of this study. To

be applicable to the later optimization problem, the metric used as a surrogate to represent this workload category is a simple, additive one, i.e. the occurrence of flight proximity. An enroute separation standard is used for defining proximity (FAA, 2010): If any two flight tracks intersect within a perimeter of 10 nautical miles horizontally and 20 Flight Levels vertically and within a 5-minute interval, this will be considered an occurrence of flight proximity in this study.

5.4.2 Model Specification and Estimation Results

The categorical response variable is the number of controllers serving a sector and defined as follows:

- NPos – the maximum number of control positions simultaneously serving a sector in a defined period.

All the traffic metrics are computed based on FPAs in the Minneapolis Center (ZMP) per defined time bin from 07/02/2007 to 07/04/2007 during the hours between 1500 and 2300 (GMT Time), and the dynamic FPA combination information is used to define the effective control region of a sector so that those previously defined metrics can be aggregated into sector-based ones. The set of covariates consists of the following:

- Hit – the number of radar hits in the effective control region of a sector in a defined period that arise from the sector's FPAs.
- Cross – the number of flights transferring from/to the effective control region of a sector in a defined period.

- Proximity– the number of flight proximities identified in the effective control region of a sector in a defined period.

The covariates and their combination effects are considered in the development of ordinal probit models. The latent function for this study is then specified as:

$$y^* = \beta_H \text{Hit} + \beta_{Cross} \text{Cross} + \beta_{prox} \text{Proximity} + \varepsilon$$

Intuitively, each sector may have individual differences and staffing conventions. Some sectors are at low altitudes and cover major airports, whereas some are high-altitude sectors and handle more overflight operations. Some sectors rarely use the second position. Thus, to increase the predictability of staffing decisions and facilitate the optimization model in the following chapter, the proposed statistical approach will be applied to each sector, so sector-specific models will be estimated.

It is also observed that the observations of multiple positions in use, especially three positions, are much fewer than those for a single position in use. The histogram of hit counts by positions in Fig 5-4 shows that the traffic intensity of using two or three positions is not distinct enough. Our computation experience also suggests that considering separately the second and third position is not statistically significant and does not help improve the model's explanatory power, so we decide to model two staffing choices, i.e. using one or more than one positions.

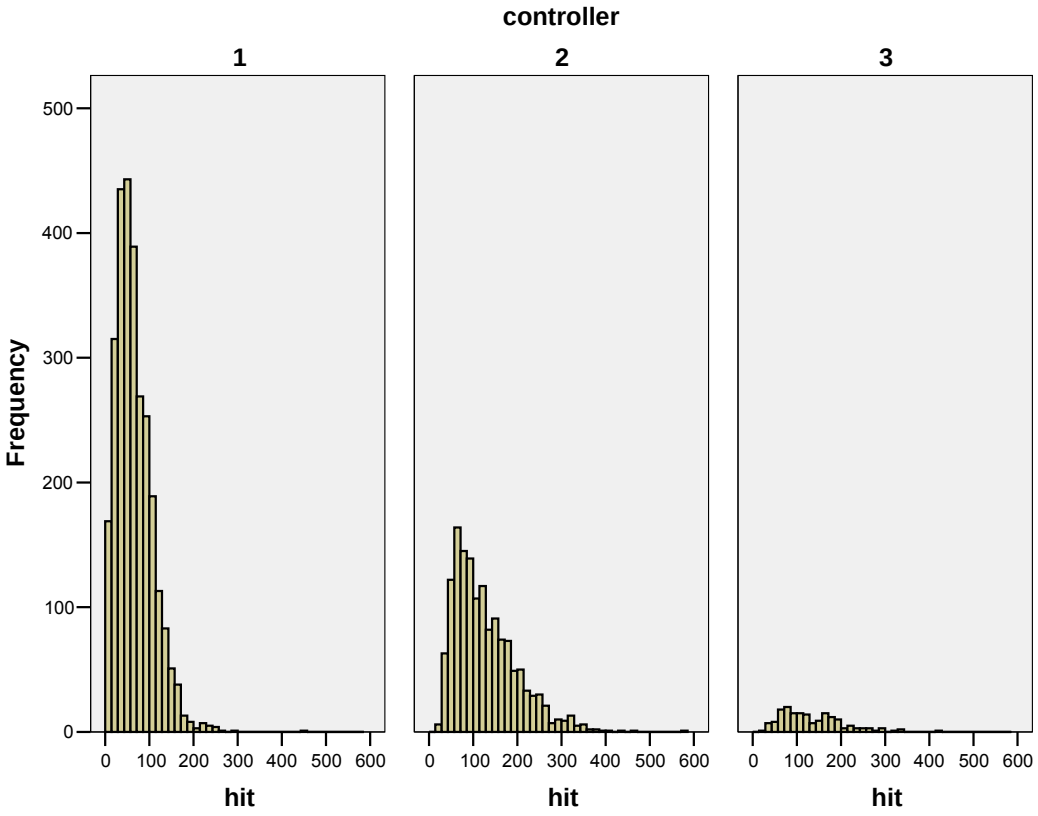


Figure 5-4 The Histograms of [Hit] vs. Controller Staffing Decisions

Finally, to investigate the effect of the defined observation period, two choices of time bin are considered, i.e. 15-minute period and 5-minute period.

In Tables 5-1 and 5-2, the estimation results of the 15-minute models and 5-minute models including the parameters of thresholds and covariates as well as basic descriptive statistics are summarized by six areas of operation[4]. For some sectors, either no or few multiple positions is observed during the selected time period, which makes the statistical estimation unnecessary.

Table 5-1 Estimation Results of the 15-Minute Models

[4] An area of operation consists of a group of sectors requiring the service of ATCSs. The number of areas authorized is based on the ARTCC's requirements and staffing needs. Vice President of En Route and Oceanic Services approval must be obtained prior to changing the number of areas of operation. (FAA, 2010)

Area of Operation	1					
Sector ID	1*	2*	3	4	12	13
Floor Altitude	0	0	0	0	240	240
Ceiling Altitude	239	239	239	239	999	999
Sample Size	41	108	58	108	108	105
P-value for LR Test	0.244	0.890	0.027	0.001	0.080	0.000
PseudoR2	0.237	0.008	0.280	0.203	0.081	0.261
[Threshold]	2.878	-0.366	3.155	1.604	0.893	2.344
[Cross]	0.049	0.007	0.047	0.127	0.092	0.052
[Conflict]	-0.046	0.047	-0.315	-0.085	0.019	-0.193
[Hit]	0.021	-0.001	0.027	0.010	0.003	0.015
P-value of [Threshold]	0.001	0.289	0.001	0.000	0.036	0.000
P-value of [Cross]	0.812	0.905	0.633	0.022	0.111	0.361
P-value of [Conflict]	0.825	0.468	0.192	0.573	0.771	0.024
P-value of [Hit]	0.214	0.818	0.010	0.219	0.301	0.001
Correct Prediction(%)	95%	65%	91%	76%	57%	70%

* Insignificant model.

Area of Operation	2					
Sector ID	5*	6	10	15	16	21*
Floor Altitude	0	80	0	240	240	80
Ceiling Altitude	239	239	239	339	339	239
Sample Size	107	95	94	79	120	64
P-value for LR Test	0.554	0.006	0.002	0.000	0.000	0.240
PseudoR2	0.026	0.188	0.196	0.508	0.649	0.170
[Threshold]	0.070	1.383	1.304	4.098	3.480	1.885
[Cross]	-0.035	0.077	0.020	-0.048	0.130	0.027
[Conflict]	0.174	0.118	0.080	-0.049	0.382	0.352
[Hit]	0.002	-0.008	0.007	0.065	-0.009	-0.004
P-value of [Threshold]	0.777	0.000	0.000	0.001	0.000	0.002
P-value of [Cross]	0.368	0.085	0.700	0.755	0.263	0.764
P-value of [Conflict]	0.257	0.021	0.433	0.798	0.036	0.124
P-value of [Hit]	0.725	0.145	0.134	0.011	0.646	0.809
Correct Prediction(%)	55%	80%	72%	94%	97%	92%

* Insignificant model.

Area of Operation	3					
Sector ID	7*	8*	9*	17*	18*	19
Floor Altitude	91	0	0	240	240	240
Ceiling Altitude	239	239	239	339	339	999
Sample Size	108	50	82	95	108	97

P-value for LR Test	0.205	0.214	0.231	0.375	0.258	0.011
PseudoR2	0.057	0.179	0.068	0.043	0.055	0.172
[Threshold]	0.133	-2.497	-0.053	0.331	1.519	1.840
[Cross]	0.046	-0.243	-0.021	0.088	0.033	0.018
[Conflict]	-0.007	-0.126	0.146	-0.007	0.034	0.042
[Hit]	-0.016	0.011	-0.003	-0.004	0.004	0.009
P-value of [Threshold]	0.654	0.002	0.863	0.387	0.002	0.000
P-value of [Cross]	0.252	0.056	0.696	0.107	0.470	0.807
P-value of [Conflict]	0.935	0.711	0.044	0.921	0.644	0.505
P-value of [Hit]	0.053	0.417	0.593	0.631	0.544	0.464
Correct Prediction(%)	60%	90%	59%	57%	77%	85%

* Insignificant model.

Area of Operation			**4**			
Sector ID	11*	20	23	24*	25	33
Floor Altitude	240	240	0	0	0	0
Ceiling Altitude	999	999	999	999	999	239
Sample Size	84	108	108	108	108	95
P-value for LR Test	0.358	0.000	0.000	0.988	0.001	0.000
PseudoR2	0.050	0.349	0.543	0.002	0.241	0.578
[Threshold]	-0.011	3.010	3.308	-0.399	3.209	3.185
[Cross]	-0.013	0.048	-0.094	0.006	0.090	-0.152
[Conflict]	0.135	-0.069	0.170	-0.036	-0.279	-0.585
[Hit]	0.001	0.018	0.023	0.000	0.015	0.029
P-value of [Threshold]	0.978	0.000	0.000	0.290	0.000	0.000
P-value of [Cross]	0.761	0.423	0.137	0.883	0.109	0.303
P-value of [Conflict]	0.157	0.297	0.240	0.753	0.116	0.321
P-value of [Hit]	0.861	0.000	0.000	0.989	0.004	0.011
Correct Prediction(%)	54%	84%	91%	67%	79%	98%

* Insignificant model.

Area of Operation			**5**			
Sector ID	26	28	29	30	36	37*
Floor Altitude	0	370	240	240	0	0
Ceiling Altitude	239	999	369	369	239	239
Sample Size	107	105	120	84	107	67
P-value for LR Test	0.000	0.000	0.070	0.083	0.008	0.461
PseudoR2	0.374	0.385	0.367	0.102	0.140	0.083
[Threshold]	1.834	2.003	4.348	0.033	0.319	0.970
[Cross]	0.156	-0.008	-0.010	0.137	-0.072	-0.164
[Conflict]	-0.139	0.032	0.123	0.030	-0.113	-0.029
[Hit]	0.012	0.017	0.017	-0.020	0.024	0.014

P-value of [Threshold]	0.000	0.002	0.005	0.943	0.310	0.041
P-value of [Cross]	0.000	0.915	0.939	0.026	0.070	0.147
P-value of [Conflict]	0.289	0.461	0.207	0.553	0.556	0.938
P-value of [Hit]	0.025	0.003	0.301	0.029	0.002	0.289
Correct Prediction(%)	71%	76%	98%	60%	61%	91%

* Insignificant model.

Area of Operation			**6**			
Sector ID	27	38*	39*	40*	42	43*
Floor Altitude	0	240	240	390	350	350
Ceiling Altitude	239	349	349	999	389	389
Sample Size	98	108	108		108	
P-value for LR Test	0.000	0.146	0.626		0.000	
PseudoR2	0.543	0.069	0.046		0.424	
[Threshold]	4.383	0.296	2.267		2.554	
[Cross]	0.054	0.012	0.007		-0.002	
[Conflict]	0.099	-0.133	-0.061		0.033	
[Hit]	0.040	0.000	0.011		0.017	
P-value of [Threshold]	0.000	0.408	0.000		0.000	
P-value of [Cross]	0.335	0.839	0.944		0.967	
P-value of [Conflict]	0.473	0.091	0.619		0.324	
P-value of [Hit]	0.001	0.979	0.260		0.003	
Correct Prediction(%)	92%	69%	94%		76%	

* Insignificant model or no enough observations.

Table 5-2 Estimation Results of the 5-Minute Models

Area of Operation			**1**			
Sector ID	1	2*	3	4	12	13
Floor Altitude	0	0	0	0	240	240
Ceiling Altitude	239	239	239	239	999	999
Sample Size	106	287	133	276	288	288
P-value for LR Test	0.006	0.211	0.002	0.000	0.000	0.000
PseudoR2	0.403	0.021	0.198	0.142	0.094	0.214
[Threshold]	3.547	-0.109	2.491	1.219	1.205	1.927
[Cross]	0.323	0.028	0.095	0.122	0.070	0.068
[Conflict]	-0.218	0.163	-0.383	-0.132	0.060	-0.166
[Hit]	0.055	-0.003	0.053	0.040	0.017	0.033
P-value of [Threshold]	0.000	0.561	0.000	0.000	0.000	0.000
P-value of [Cross]	0.075	0.633	0.360	0.036	0.246	0.184
P-value of [Conflict]	0.493	0.064	0.207	0.527	0.425	0.087
P-value of [Hit]	0.047	0.728	0.001	0.000	0.001	0.000

Correct Prediction(%)	98%	55%	87%	70%	63%	75%

* Insignificant model.

Area of Operation	2					
Sector ID	5*	6	10	15	16	21
Floor Altitude	0	80	0	240	240	80
Ceiling Altitude	239	239	239	339	339	239
Sample Size	277	261	275	193	338	150
P-value for LR Test	0.795	0.000	0.000	0.000	0.000	0.079
PseudoR2	0.005	0.129	0.155	0.262	0.572	0.114
[Threshold]	0.140	1.524	1.188	2.834	3.938	1.830
[Cross]	0.005	0.114	0.011	0.010	0.211	0.088
[Conflict]	0.154	0.170	0.136	0.138	0.329	0.524
[Hit]	0.001	-0.003	0.020	0.076	0.049	-0.009
P-value of [Threshold]	0.301	0.000	0.000	0.000	0.000	0.000
P-value of [Cross]	0.911	0.030	0.834	0.947	0.047	0.433
P-value of [Conflict]	0.378	0.022	0.197	0.440	0.004	0.034
P-value of [Hit]	0.890	0.621	0.001	0.009	0.019	0.642
Correct Prediction(%)	55%	82%	71%	93%	96%	93%

* Insignificant model.

Area of Operation	3					
Sector ID	7	8	9	17	18	19
Floor Altitude	91	0	0	240	240	240
Ceiling Altitude	239	239	239	339	339	999
Sample Size	300	141	305	225	324	294
P-value for LR Test	0.015	0.078	0.021	0.051	0.035	0.000
PseudoR2	0.048	0.073	0.043	0.045	0.042	0.212
[Threshold]	0.125	-0.903	0.389	0.579	1.466	2.074
[Cross]	0.040	-0.172	0.031	0.078	0.053	0.029
[Conflict]	0.027	-0.542	0.186	0.032	-0.036	0.093
[Hit]	-0.031	0.021	-0.004	0.012	0.018	0.034
P-value of [Threshold]	0.425	0.000	0.005	0.007	0.000	0.000
P-value of [Cross]	0.376	0.074	0.511	0.160	0.226	0.611
P-value of [Conflict]	0.782	0.036	0.007	0.745	0.658	0.184
P-value of [Hit]	0.002	0.104	0.464	0.221	0.051	0.002
Correct Prediction(%)	64%	79%	64%	54%	81%	85%

* Insignificant model.

Area of Operation	4					
Sector ID	11*	20	23	24*	25	33

Floor Altitude	240	240	0	0	0	0
Ceiling Altitude	999	999	999	999	999	239
Sample Size	322	324	324	324	288	280
P-value for LR Test	0.148	0.000	0.000	0.813	0.000	0.000
PseudoR2	0.022	0.316	0.489	0.004	0.150	0.676
[Threshold]	0.251	2.639	3.492	-0.107	2.293	3.659
[Cross]	-0.003	0.075	-0.005	0.028	0.053	-0.192
[Conflict]	0.102	-0.008	0.122	0.001	-0.260	-0.550
[Hit]	0.007	0.040	0.056	0.001	0.032	0.070
P-value of [Threshold]	0.141	0.000	0.000	0.574	0.000	0.000
P-value of [Cross]	0.946	0.191	0.929	0.418	0.340	0.304
P-value of [Conflict]	0.267	0.911	0.400	0.996	0.136	0.387
P-value of [Hit]	0.235	0.000	0.000	0.893	0.000	0.000
Correct Prediction(%)	52%	84%	90%	40%	82%	99%

* Insignificant model.

Area of Operation			**5**			
Sector ID	26	28	29	30	36	37*
Floor Altitude	0	370	240	240	0	0
Ceiling Altitude	239	999	369	369	239	239
Sample Size	250	252	322	251	242	186
P-value for LR Test	0.000	0.000	0.002	0.080	0.004	0.812
PseudoR2	0.289	0.321	0.274	0.036	0.070	0.016
[Threshold]	1.378	1.613	3.726	0.071	0.596	1.660
[Cross]	0.133	-0.006	0.000	0.092	-0.043	-0.097
[Conflict]	-0.209	0.007	0.059	0.056	0.066	0.117
[Hit]	0.055	0.044	0.053	-0.022	0.038	0.015
P-value of [Threshold]	0.000	0.000	0.000	0.740	0.002	0.000
P-value of [Cross]	0.008	0.935	0.998	0.074	0.379	0.426
P-value of [Conflict]	0.179	0.887	0.620	0.322	0.753	0.822
P-value of [Hit]	0.000	0.000	0.013	0.018	0.001	0.503
Correct Prediction(%)	70%	73%	98%	61%	60%	95%

* Insignificant model.

Area of Operation			**6**			
Sector ID	27	38*	39	40*	42	43
Floor Altitude	0	240	240	390	350	350
Ceiling Altitude	239	349	349	999	389	389
Sample Size	232	251	249		252	288
P-value for LR Test	0.000	0.349	0.000		0.000	0.000
PseudoR2	0.346	0.019	0.594		0.407	0.530
[Threshold]	2.873	0.382	7.016		2.568	4.791

[Cross]	0.054	0.005	0.303	0.111	0.137
[Conflict]	0.100	-0.089	-0.166	0.056	0.201
[Hit]	0.071	-0.008	0.116	0.041	0.032
P-value of [Threshold]	0.000	0.049	0.002	0.000	0.000
P-value of [Cross]	0.357	0.929	0.263	0.066	0.259
P-value of [Conflict]	0.480	0.314	0.497	0.173	0.013
P-value of [Hit]	0.000	0.505	0.013	0.000	0.357
Correct Prediction(%)	88%	72%	99%	73%	99%

***** Insignificant model or no enough observations.

5.5 Result Interpretations

In Tables 5-1 and 5-2, the likelihood ratio (LR) test is intended to examine the improvement achieved by the model using specified covariates besides the intercept.

$$\text{LR} = -2\frac{\ln L(M_{Intercept})}{\ln L(M_{Specified})} \sim \chi_v^2$$

where $\nu = 3$ for this study since three covariates are specified.

If the chi-square value of the LR test is less or equal than the significance level, it suggests the rejection of the null hypothesis that the specified model has no difference from the intercept only model, which implies the specification is statistically significant.

The pseudo r-squared values can be used to compare the goodness-of-fit among models rather than measure models' prediction power of categorical responses. An issue of interpreting this value is that except for 0 or 1, the pseudo r-square value does not have as intuitive explanation as the r-square value for linear regression model (Long, 1997). While there are several definitions of pseudo r-square to assess the goodness-of fit of the ordinal regression model, we use the measure developed by Nagelkerke (1991):

$$\text{Nagelkerke's Pseudo R-Squared} = \frac{1-\left(\frac{L(M_{Intercept})}{L(M_{Specified})}\right)^{2/N}}{1-L(M_{Intercept})^{2/N}}$$

where N is the sample size.

In addition to the LR ratio and pseudo R-squared, we also compute the percentage of correct predictions by comparing the actual and predicted staffing categories. One should note that the ordinal regression model attempts to predict cumulative probabilities rather than staffing decisions, which suggests the pseudo r-square should not be the only indicator of how the estimated models perform. Since our goal is to use the models to replicate sector staffing decisions, the correct prediction rate is also an important metric for judging the predictive ability of the models. Two steps are required to get predicted staffing categories. First, for each observation, the probabilities must be estimated for each category. Second, those probabilities must be used to select the most likely outcome category for each observation. The predicted category, i.e. the number of controllers used, is then identified as the one with the highest probability.

If we set 0.1 as the significance level for LR Test, there are 7 out of 36 sectors showing insignificance of the models among the 5-minute models, whereas there are 16 sectors showing insignificance of the models among the 15-minute models. There are 25 and 22 sectors having correct prediction rate over 70% among the 5- and 15-minute models, respectively.

The effect of the covariates to the probabilities of staffing decisions is nonlinear. Fig 5-5 shows the effect of increases in a covariate with positive coefficient on probabilities of estimated sector staffing. When the value of covariate increases, the

mean of the distribution moves rightward, meaning a decrease of probability of using 1 position. As the mean passes the estimated threshold value (the black vertical line), the category of using 2+ positions has higher probability and is then the category predicted by the model. For example, Sector 27 in Area 6 of the 15-min results has positive coefficient of [Hit]. By varying the values of [Hit] and holding other covariates at their medians, Fig 5-6 shows the changes in the probability of using 1 and 2+ positions. Before the number of hits reaches 95, the most likely outcome is using 1 position. The probability of using 2+ positions is actually the complement of using 1 position.

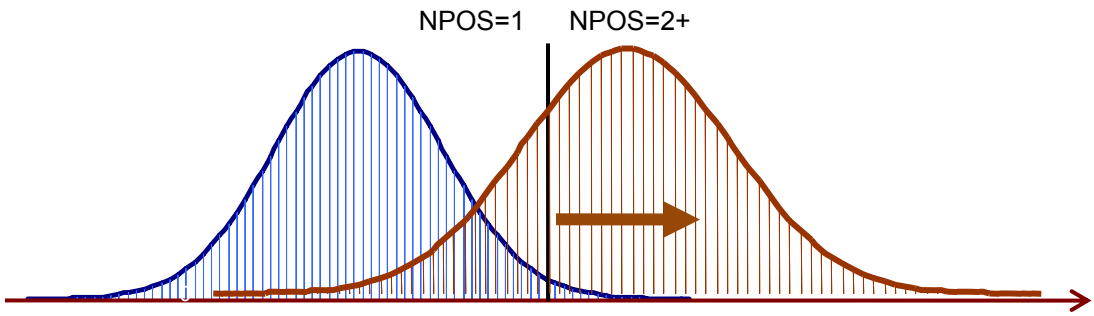


Figure 5-5 Effect of increases in a covariate on probabilities of estimated sector staffing if $\beta>0$

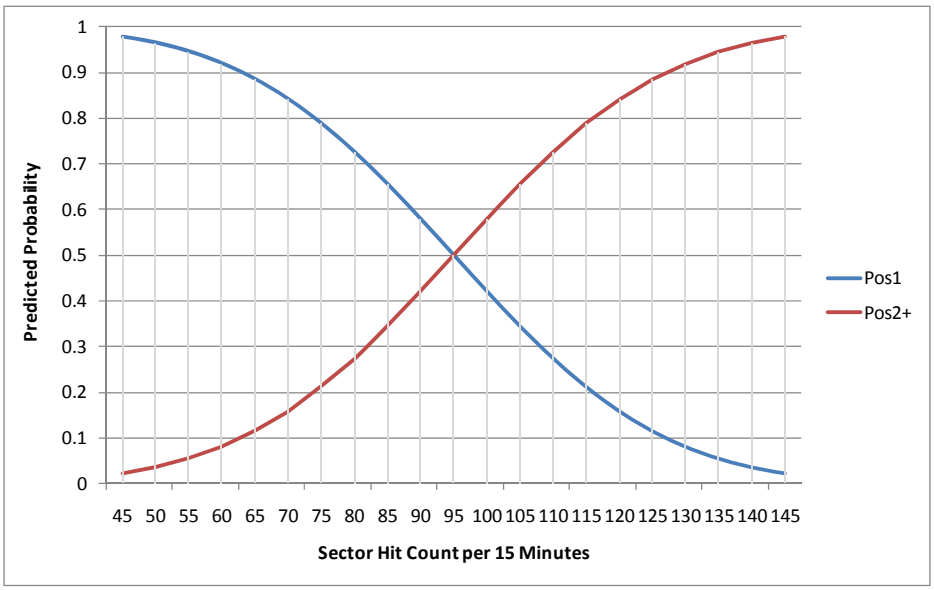


Figure 5-6 Sensitivity of Hit Counts on Sector 27 from the Results of the15-Minute Models

Fig 5-7 shows the effect of the interaction between two covariates. The curves represent the probability changes of using 1 position at the 5th, 25th, 50th, 75th, and 95th percentiles of covariate [Cross]. Since the coefficient of [Cross] is positive for Sector 27, with the same level of [Hit] the higher the [Cross] the lower is the probability of using 1 position. Note that the gaps between two adjacent curves are not equally spaced, demonstrating a nonlinear relation between the probability and the crossing traffic in Sector 27.

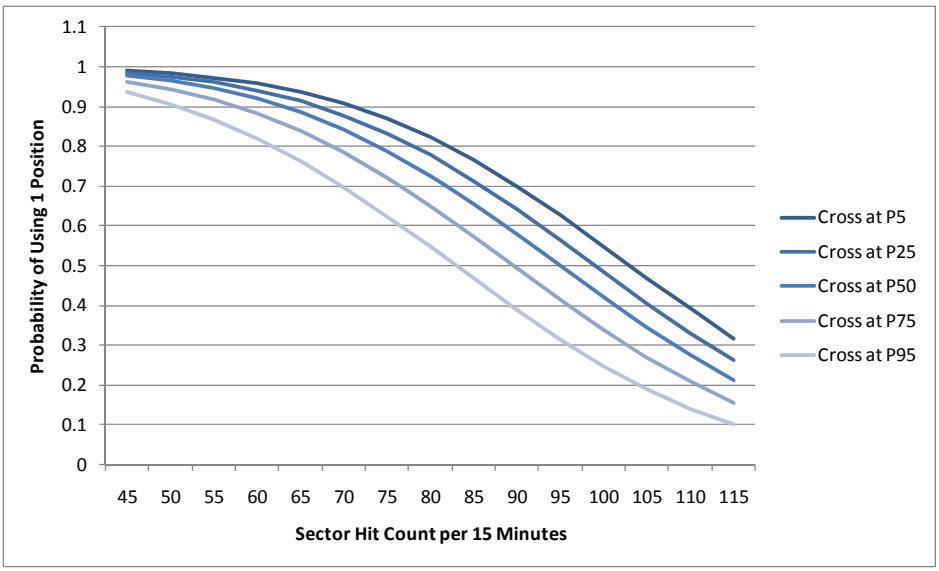
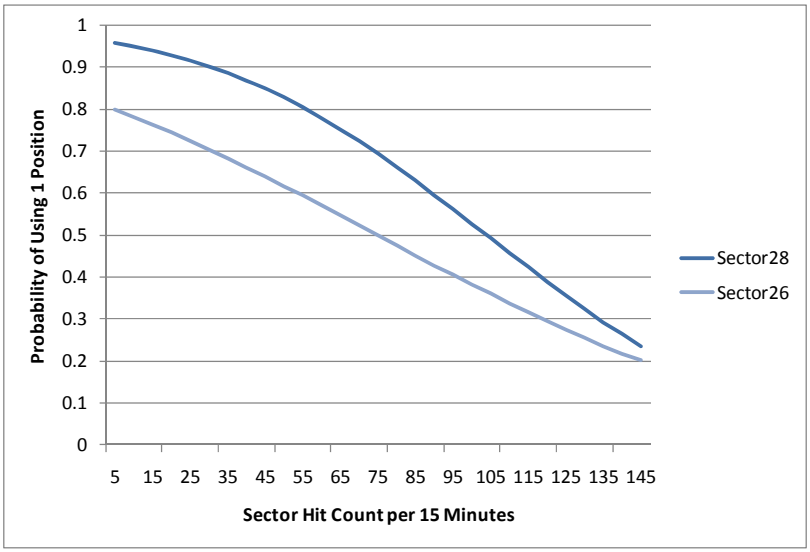


Figure 5-7 Sensitivity to Hit Counts and Flight Crossings in Sector 27 from the Results of the 15-Minute Models

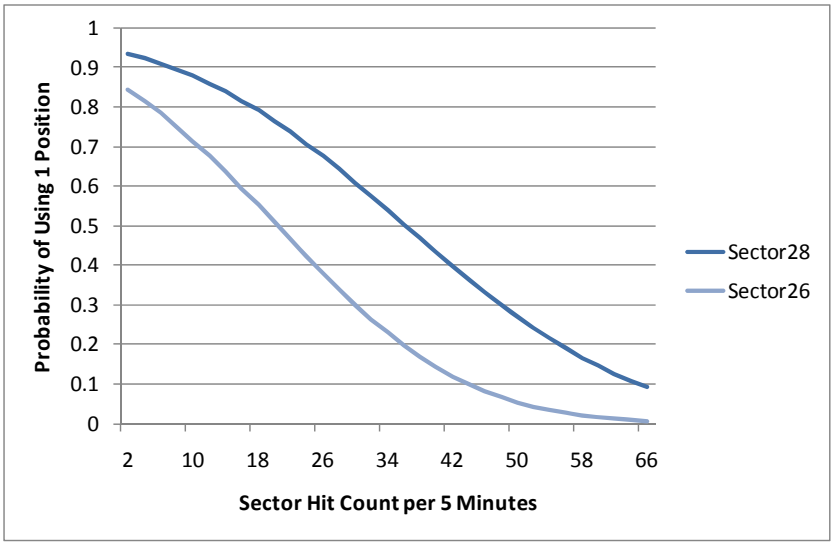
Another dimension of interest worth examining is the duration of observation periods, i.e. 15-minute or 5 minute. Sectors 26 and 28 in Area 5 are used for illustration. Fig 5-8(a) shows the results of the 15-minute models, presenting different behaviors of probability curves for two sectors. With all other covariates held at their medians, the turning point of using 2+ positions in Sector 26 is 30 hits below that in Sector 28. The horizontal axis is bounded by the range of [Hit] of these two sectors. It is apparent that Sector 26 demonstrates a linear relation between the covariate [Hit] and the probability of using 1 position, whereas Sector 28 has a nonlinear one.

Similar interpretations apply to Fig 5-8(b), which illustrates the 5-minute results. However, the probability curve of Sector 26 shows a nonlinear trend. The probability drops faster than that from the 15-minute results as the value of covariate [Hit] increases. For Sector 28, the probability also drops drastically within a relatively shorter range of [Hit] than that from the 15-minute results. This implies that when the observation period is shorter, the probability of using 2+ position is more sensitive to

the covariate [Hit]. If the traffic increases in a relatively short time, the controller workload increases drastically and results in higher probability of using multiple positions.



(a) Results of the15-Minute Models



(b) Results of the 5-Minute Models

Figure 5-8 Sensitivity to Hit Counts in Sectors 26 and 28

5.6 Summary

In this chapter, we have applied ordinal regression analysis on sectors of the Minneapolis enroute center and estimated a set of statistical models for individual sectors for predicting staffing decisions. Specifically, the probability of each staffing category can be computed by using selected metrics and associated coefficients from the estimation results. The category with the highest probability is then the predicted staffing decision.

The results in this chapter extend the previously assumed relation that controller staffing is a deterministic, step-wise function of traffic into a probabilistic, nonlinear one. This relaxation is meaningful when the length of each design period is shortened from 2 hours to 15 minutes. Comparing the results from the models with different period durations, it is also found that the nonlinearity is more significant in the 5-minute models than in the 15-minute ones (as illustrated by Sector 26).

Conceptually, during a short response time, 15 or even 5 minutes, the controller staffing cannot easily follow a pre-specified staffing standards (such as Fig. 1-3), especially in the heat of the battles. The area supervisors will use their own judgments in assigning control positions. The threshold for adding another controller to help with traffic should no longer be treated as a fixed, single point on the scale of traffic complexity but a range with high likelihood of changing staffing decisions. The observations and analyses made in this chapter confirm this concept and can be further incorporated into designing an operational model for the Sector Combination Problem.

Chapter 6: Sector Combination Problem with Consideration of Staffing Efficiency

6.1 Introduction

For day-to-day operations, a common way to deal with traffic demand variation over time and space is to temporarily combine sectors with low traffic or assign more controllers to busy sectors. For staff and workload management purposes, enroute sectors are grouped into "areas" of operations, which typically comprise six to eight sectors. Currently, sector combinations are limited to areas, and can occur as frequently as every half hour with as little lead time as a few minutes. Area supervisors make these decisions based upon their experience and judgments.

Bloem et al. (2009) summarizes sector-combining issues and procedures. They state that in the current ATC environment, stability and controllers' familiarity of the sector combinations would be major concerns of the area supervisors. On the other hand, there are the benefits expected from sector combination/splitting activities. Combining sectors reduces controller staff required to manage a volume of airspace and leads to fewer airspace-induced flight restrictions, e.g. more direct routings. When splitting sectors, greater traffic demand can be accommodated. When traffic is lighter, controllers can satisfy more pilot requests such as altitude changes to avoid turbulence.

Combining sectors in response to traffic demand changes promotes safety and efficient use of air traffic control resources (e.g. controllers). Highly frequent changes can become impractical for controllers to manage. Advancements in NextGen

technologies (e.g. data link, automated conflict detection with proposed resolutions, ADS-B) promise to alleviate some of this burden (Gupta et al., 2009).

Research on combining sectors to achieve system efficiency has been carried out for years, but none of the existing studies incorporates staffing concerns. Among European studies, Delahaye et al. (1995) apply a genetic algorithm to find the combination that balances the workload among combined sectors and assume that the capacity of a combined sector is the maximum value among the sectors before being combined. Verlhac and Machon (2001) propose integer programming models that search for a suitable set of pre-defined layouts to best accommodate time-varying demand by minimizing sector capacity deficits. Gianazza, et al. (2002a) and Gianazza and Alliot (2002b) propose a cost function for optimizing sector combination by weighting two design objectives, i.e. the number of resulting sector control positions and the deficit and surplus of sector capacity. Later, by training an artificial neural network model with historical data, Gianazza (2007, 2008, 2009) use realistic measures of traffic complexity to predict the probability of sectors being merged, manned, or split. Then, potential airspace configurations, no longer tied to a set of pre-defined ones, are enumerated and evaluated to find the best configuration. Recently, the sector combination problem has received attention from U.S. researchers. Bloem et al. (2009) develop a local improvement heuristic to combine adjacent under-utilized sectors and to minimize the number of sectors after combination, and then Drew (2009) develops an optimization version of Bloem et al's problem by formulating a variant of the multi-commodity network flow problem. Klein et al. (2008) uses FPAs as fundamental blocks for defining sector boundaries

and balancing sector workload, which does not guarantee staffing efficiency. Although the approaches applied in the literature are different, a shared goal is to minimize the sector count.

In this chapter, an optimization model is proposed that incorporates controller staffing concerns into sector combination decisions. It is supposed to help the decision making of area supervisors on how the sectors should be combined and what level of controller staffing is required. In particular, the fixed posting areas (FPAs)[5] are used as the fundamental units for combining/splitting sectors. As sector volume usually consists of several FPAs, the combination/split activities based on FPAs are commonly observed. If a sector is overloaded, then in addition to adding more controllers to serve that sector, a practical approach is to assign one or more of its (FPAs) to its adjacent sector(s). Moreover, the proposed model proposed differs from any existing study in two ways: First, instead of using deterministic staffing standards, a statistical model is estimated to predict staffing decisions based on traffic characteristics. Second, a novel mixed integer program is formulated that incorporates the statistical results and finds the sector combination with minimal staffing level. It extends the research scope of existing studies and helps determine the efficient controller staffing.

The chapter is organized as follows. Section 6.2 formally defines the problem of sector/FPA combination and controller staffing and formulates a mathematical program. In Section 6.3, appropriate approximation or linearization techniques are

[5] A fixed posting area (FPA) is a three-dimensional volume of airspace and can be considered as a fundamental unit of airspace. The airspace of a sector is a set of one or more contiguous FPAs that constitute a specified sector. An FPA has a default sector but may be designated to others due to ATC operational needs.

applied to transform the proposed formulation into a linear mixed-integer program (MIP). A customized algorithm is developed so that the proposed MIP can be solved with commercial solvers. Numerical experiments are conducted and their results are analyzed in Section 6.4. To address the issue of combination stability over time, which is the main concern of practitioners, a time-dependent version of the sector/FPA combination problem is proposed. Concluding remarks follow in Section 6.5.

6.2 Optimizing Sector/FPA Combination Schemes

6.2.1 Problem Statement

In this study, the Sector/FPA Combination Problem is defined as deciding how the sectors should be staffed and FPAs should be combined or split, given the traffic forecast over time and space. Its main purpose is to support the decision making process of air traffic managers/supervisors at enroute centers who encounter this problem or similar types on a daily basis. Depending on how often the problem is considered, the resource of interest is the total controller shifts or hours. Thus, the objective of this problem is to find an efficient controller staffing plan for serving sector traffic via optimizing FPA combination strategies.

An intuitive thought for dealing with this problem is to directly apply the proposed sectorization models from Chapter 3. However, under a more dynamic environment (e.g. every 15-min period), the relation between workload metrics and staffing decisions should no longer be assumed to be linear. The statistical results from the previous section on the relation between staffing decisions and traffic metrics are

expected to be incorporated into the proposed optimization model, so a major modification to address additional concerns of this problem is necessary.

The problem also extends the scope of prior studies by minimizing not only sector hours but also controller resources. Note that efficient controller staffing does not mean compromising safety or having negative impact of labor contracts. Instead, it should be considered as providing the managerial roles a baseline for assigning controller workforce in a more flexible way and also advising the managers a minimal workforce required to maintain a safe operation. Especially, under a facility-free ATC environment envisioned in NextGen, controllers might be able to work on any airspace from wherever they are, so controller resource management would be more flexible than today.

6.2.2 Mathematical Formulation

In this section, a mathematical formulation is proposed to tackle the sector/FPA combination problem considering staffing efficiency. Denote each FPA with index $i \in \{1,...,I\}$. Given the adjacency relations of sectors and their FPAs in the target airspace, edge set $E = \{(i,j) \mid i = 1,...,I, j \in \delta(i), i > j\}$ describe the geographical adjacency relation of FPAs. Each FPA can thus be represented by a node in a network, where there is an edge connecting two adjacent FPAs in such a network.

Sector Connectivity

The basic requirement for bringing FPAs into sectors is to maintain the connectivity of a sector, i.e. all FPAs assigned to a sector should be geographically connected. Denoting sector index as $k \in \{1,...,K\}$, the decision variables and the constraints that govern sector connectivity are defined and formulated as follows:

$$x_{i,k} = \begin{cases} 1, & \text{if node } i \text{ is assigned to sector } k. \\ 0, & \text{otherwise.} \end{cases}$$

$$y_{ij,k} = \begin{cases} 1, & \text{if edge}(i, j) \text{ is assigned to sector } k. \\ 0, & \text{otherwise.} \end{cases}$$

$$z_k = \begin{cases} 1, & \text{if sector } k \text{ is employed to serve traffic.} \\ 0, & \text{otherwise.} \end{cases}$$

FPA assignment constraint:

$$\sum_k x_{i,k} = 1 \quad \text{for all } i \in \{1,...,I\} \tag{6-1}$$

Edge assignment constraints:

$$y_{ij,k} \le \begin{cases} x_{i,k} \\ x_{j,k} \end{cases} \quad \text{for all } (i, j) \in E, k \in \{1,...,K\} \tag{6-2}$$

Connectivity constraints:

$$x_{i,k} \le z_k \quad \text{for all } i \in \{1,...,I\},\ k \in \{1,...,K\} \tag{6-3}$$

$$\sum_{(i,j)\in E} y_{ij,k} = \sum_i x_{i,k} - z_k \quad \text{for all } k \in \{1,...,K\} \tag{6-4}$$

$$\sum_{(i,j)\in E(X)} y_{ij,k} \le |X| - 1 \quad \text{for any } X \subset \{1,...,I\}, |X| \ge 2, \text{and} \quad \text{for all } k \in \{1,...,K\} \tag{6-5}$$

Constraint (6-1) limits each node to be assigned to only one sector. Constraint (6-2) ensures that if two adjacent FPAs, i and j, are not assigned to a sector, edge (i, j) will not be assigned to that sector.

Constraint (6-3) says if $x_{i,k} = 1$ for any i, then sector k is active, i.e. $z_k = 1$. Constraints (6-4) and (6-5) are adopted from the formulation of the spanning tree polytope (Edmond, 1970), maintaining a tree structure for the edges assigned to

individual sector in order to ensure the connectivity of the FPAs assigned in the same sector. Constraint (6-4) is modified from the cardinality constraint: for all $k \in \{1,...,K\}$, if sector k is employed and assigned positive demand (or workload) of connected FPAs, i.e. $z_k = 1$, the number of edges assigned to sector k is equal to the number of FPAs (or nodes) assigned sector k less one. Otherwise, both sides of the equation are thus zero, implying sector k is not active. Constraint (6-5) is referred as a ***Cycle Elimination Constraint***: for any subset of FPAs (or nodes) assigned to a sector, the number of edges assigned to a sector cannot exceed the cardinality less one. The number of constraints of this type grows exponentially with the number of FPAs. For a realistic problem size, Constraint (7) cannot be enumerated exhaustively a priori, so a special treatment is needed.

Workload Measurement and Aggregation

Controller workload is further decomposed into three categories used in Delahaye (1995). Denote $w^m{}_i$ and $w^f{}_i$ the controller workload of traffic monitoring and conflict resolving, respectively, known (forecasted) in FPA i and $w^c{}_{ij}$ the workload of coordination between FPAs i and j. Let decision variables $\overline{w^m}_k$, $\overline{w^f}_k$, $\overline{w^c}_k$ represent the aggregated workload of traffic monitoring, conflict resolving, and coordination, respectively, for sector k. The advantage of separately considering workload in FPA into three categories is that a more detailed treatment of aggregating FPA workload metrics into sector-based one becomes possible. For example, when two FPAs are combined, the coordination workload between them will no longer be

counted. This then relaxes the assumption in the sectorization model that all workload metrics are additive.

Workload aggregation constraints:

$$\sum_i w^m{}_i x_{i,k} = \overline{w^m}_k \quad \text{for all } k \in \{1,...,K\} \tag{6-6}$$

$$\sum_i w^f{}_i x_{i,k} = \overline{w^f}_k \quad \text{for all } k \in \{1,...,K\} \tag{6-7}$$

$$\sum_{(i,i)\in E} w^c{}_{ij} \left| x_{i,k} - x_{j,k} \right| = \overline{w^c}_k \quad \text{for all } k \in \{1,...,K\} \tag{6-8}$$

Constraints (6-6), (6-7), and (6-8) are workload aggregation constraints. With them, the resulting workload of each category of a sector can be computed. For coordination workload, a special treatment is made. Since the coordination workload is measured between two adjacent FPAs and associated with the edge that connects them, Constraint (6-8) states that the coordination workload aggregated from FPAs to a sector will only count the edges connecting to another sector. Thus, for each edge $(i, j) \in E$, if either i or j is assigned to sector k and the other is not, then the edge weight $w^c{}_{ij}$ will be counted as part of sector k's coordination workload.

Incorporating Controller Staffing Models into the Objective Functions

So far the constraints of setting sector capacity have not been introduced. Instead of setting a deterministic value as the sector capacity associated with each staffing decision, the statistical results are used to predict the probabilities of staffing decisions, given aggregate sector workload metrics, so the number of controllers serving sector traffic would become probabilistic. With the objective function set to minimize the predicted or expected controller usage, an optimal FPA combination strategy may be found.

A variable $\mathrm{P}_k(n)$ is defined as the probability of using n controllers on sector k to be computed using aggregated sector workload metrics, where $n = 1,...,N$. Two alternate objectives on controller costs will be examined:

1) Minimize the total predicted controller shifts

$$\min \sum_k \arg\max_n \left\{ \mathrm{P}_{k,n} \middle| n = 1,..,N \right\} \tag{6-9}$$

2) Minimize the total expected controller shifts

$$\min \sum_k \sum_n n \cdot \mathrm{P}_{k,n} \tag{6-10}$$

where

- $\mathrm{P}_{k,n=1} = \Phi(\mu_{k,n=1} - \mathbf{W}_k \boldsymbol{\beta}_k)$;
- $\mathrm{P}_{k,n=2} = \Phi(\mu_{k,n=2} - \mathbf{W}_k \boldsymbol{\beta}_k) - \Phi(\mu_{k,n=1} - \mathbf{W}_k \boldsymbol{\beta}_k)$;
- $\vdots$
- $\mathrm{P}_{k,n=N} = 1 - \Phi(\mu_{k,n=N-1} - \mathbf{W}_k \boldsymbol{\beta}_k)$;
- $\mathbf{W}_k = \left[\overline{w^m}_k, \overline{w^f}_k, \overline{w^c}_k \right]$
- $\boldsymbol{\beta}_k$ and $\mu_{k,n}$,where $n = 1,...,N-1$, are the coefficient vector and threshold values for ordinal categories from the statistical analyses.
- $\Phi(\cdot)$ is the cumulative distribution function (CDF) of the standard normal distribution.

By doing so, the staffing decision is no longer treated as deterministic but probabilistic. It is observed from the staffing records and historical data that the staffing decisions do not show a consistent pattern, even when the traffic conditions

are similar. Not surprisingly, traffic would not be the only factor considered when deciding staffing level. Controllers' experience levels, redundancy for safety, and facility culture will also impact such decisions. Since the relation between traffic and staffing decisions has been quantified, the staffing decisions can be mimicked by predicting the probability of each staffing decision through incorporating the sector-specific results for ordinal regression analyses.

One caveat is that there is no limit set for the covariate value $\mathbf{W}_k\boldsymbol{\beta}_k$, so the probability of the last ordinal category (i.e. $n = N$) would not exceed 1 even when $\mathbf{W}_k\boldsymbol{\beta}_k$ goes to infinity. To fix this condition, a set of upper bound constraints can be considered. In addition, when there no FPA assigned to a sector k, i.e. $\mathbf{W}_k\boldsymbol{\beta}_k = 0$, each staffing decision still has a predicted probability, which should not be counted in the objective function values. Modification of the objective functions can offset this concern. The fixing procedures will be further discussed in the numerical experiment section.

With the constraints (6-1)–(6-8) defined and the alternate objective functions described (6-9) and (6-10), the optimization model for Sector Combination Problem is formulated. It has an exponential number of constraints as well as non-linear objective functions. Either metaheuristics, such as a genetic algorithm, or a numerical approximation method may be suitable. Since the solution technique has been developed for the sectorization problem in Chapter 3, a straightforward approach is to apply the branch-and-cut algorithm in Chapter 3.4. In order to be handled by a MIP solver, linearization techniques must be applied to the two alternate non-linear objective functions, as explained in the following section.

6.3 Model Computability and Linear Approximation

6.3.1 Piece-wise Approximation of Probability Function

To compute $P_{k,n}$, we first need a piece-wise linear approximation of the standard normal cumulative distribution function:

$$\Phi(\varepsilon) = \int_{-\infty}^{\varepsilon} \frac{1}{\sqrt{2\pi}} \exp(\frac{\nu^2}{-2}) d\nu$$

In Table 6-1, there are 8 grid points selected to discretize the CDF, which divide the domain $[\varepsilon_1, \varepsilon_8]$ into 7 intervals. Each interval of the CDF can thus be represented by a linear function. Fig 6-1 shows the grid points and the linearized normal CDF.

Table 6-1 Grid Points to Approximate the Normal CDF

Grid Points	*l*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
Values	ε_l	-100	-4	-2	-1	1	2	4	100
Normal CDF Values	$\Phi(\varepsilon_l)$	0.00000	0.00003	0.02275	0.15866	0.84134	0.97725	0.99997	1.00000

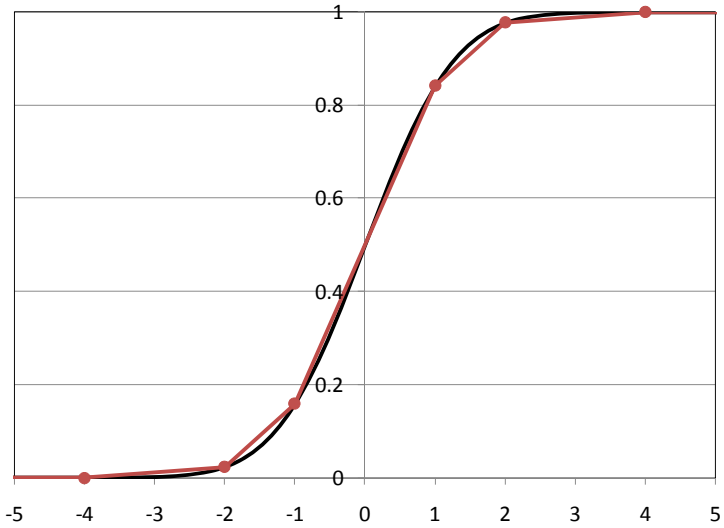


Figure 6-1 Piece-wise Linearization of Normal CDF

The next step is to determine which interval is chosen for the value to be evaluated and to compute the approximated CDF values. The approximation method proposed in Babayev (1997) is thus adopted for this study. For $k \in \{1,...,K\}$, $n \in \{1,...,N-1\}$, a set of decision variables is introduced:

- $\omega_{k,n,l} \in \{0,1\}$ for each interval $[\varepsilon_l, \varepsilon_{l+1}]$, where $l \in \{1,...,7\}$ to select one of the intervals in the linearized CDF, and
- $\lambda_{k,n,l} \geq 0$ for grid point $l \in \{1,...,8\}$ to describe the convex combination of the values of the end points on the selected interval.
- $\phi_{k,n} \geq 0$ as the approximate value for the evaluation of $\Phi(\mu_{k,n} - \boldsymbol{\beta}_k' \mathbf{W}_k)$.

To compute the approximated normal CDF values $\phi_{k,n}$ for $n \in \{1,...,N-1\}$ as well as the probability $\mathrm{P}_{k,n}$, the following constraints are added:

$$\mu_{k,n} - \boldsymbol{\beta}_k' \mathbf{W}_k = \sum_{l=1}^{8} \lambda_{k,n,l} \cdot \varepsilon_l \quad \text{for } k \in \{1,...,K\},\ n \in \{1,...,N-1\} \tag{6-11}$$

$$\phi_{k,n} = \sum_{l=1}^{8} \lambda_{k,n,l} \cdot \Phi(\varepsilon_l) \quad \text{for } k \in \{1,...,K\},\ n \in \{1,...,N-1\} \tag{6-12}$$

$$\sum_{l=1}^{8} \lambda_{k,n,l} = 1 \quad \text{for } k \in \{1,...,K\},\ n \in \{1,...,N-1\} \tag{6-13}$$

$$\sum_{l=1}^{7} \omega_{k,n,l} = 1 \quad \text{for } k \in \{1,...,K\},\ n \in \{1,...,N-1\} \tag{6-14}$$

$$\lambda_{k,n,l} \leq \omega_{k,n,l} + \omega_{k,n,l-1} \quad \text{for } k \in \{1,...,K\},\ n \in \{1,...,N-1\},\ l = \{2,...,7\} \tag{6-15a}$$

$$\lambda_{k,n,l} \leq \omega_{k,n,l} \quad \text{for } k \in \{1,...,K\},\ n \in \{1,...,N-1\},\ l = 1 \tag{6-15b}$$

$$\lambda_{k,n,l} \leq \omega_{k,n,l-1} \quad \text{for } k \in \{1,...,K\},\ n \in \{1,...,N-1\},\ l = 8 \tag{6-15c}$$

$$\mathrm{P}_{k,n} = \phi_{k,n} - \phi_{k,n-1} \quad \text{for } k \in \{1,...,K\},\ n \in \{2,...,N-1\} \tag{6-16a}$$

$$\mathrm{P}_{k,n} = \phi_{k,n} \quad \text{for } k \in \{1,...,K\},\ n = 1 \tag{6-16b}$$

$$\mathrm{P}_{k,n} = 1 - \phi_{k,n-1} \quad \text{for } k \in \{1,...,K\},\ n = N \tag{6-16c}$$

Constraints (6-11), (6-12), and (6-13) are needed to describe the argument for the normal CDF $\Phi(\bullet)$ and its approximate outcome after the evaluation of $\Phi(\bullet)$ as a convex combination of the grid points ε_l and their associated $\Phi(\varepsilon_l)$. According to Constraints (6-14) and (6-15) only one interval will be selected in this piece-wise approximation. At most two $\lambda_{k,n,l}$ may get a positive value, and in the case of two, they will be adjacent. Constraint (6-16) thus uses the approximated values to compute the staffing probabilities.

6.3.2 Transformation of the argmax Function

To transform the objective function (6-9) into a linear form, a variable that represents the predicted staffing level is introduced:

$$p_{k,n} = \begin{cases} 1, \text{ if the use of } n \text{ controller(s) on sector } k \text{ has the highest probability.} \\ 0, \text{ otherwise.} \end{cases}$$

The following constraints will be added to transform $\arg\max_n \left\{ \mathrm{P}_{k,n} \middle| n = 1,...,N \right\}$:

$$q_k \geq \mathrm{P}_{k,n} \quad \text{for all } k \in \{1,...,K\},\ n \in \{1,...,N\} \tag{6-18}$$

$$1 - p_{k,n} \geq q_k - \mathrm{P}_{k,n} \quad \text{for all } k \in \{1,...,K\},\ n \in \{1,...,N\} \tag{6-19}$$

$$\sum_n p_{k,n} = 1 \quad \text{for all } k \in \{1,...,K\} \tag{6-20}$$

A variable $q_k \geq 0$ for all $k \in \{1,...,K\}$ is used to represent the highest value among $\mathrm{P}_{k,n}$, where $n \in \{1,...,N\}$. The constrains (6-18), (6-19) and (6-20) together find the index n which has the maximum of $\mathrm{P}_{k,n}$ by letting $p_{k,n} = 1$. This works if $\mathrm{P}_{k,n} \leq 1$ for all $n \in \{1,...,N\}$. Thus, minimizing $\arg\max_n \left\{ \mathrm{P}_{k,n} \middle| n = 1,...,N \right\}$ is equivalent to minimizing $\sum_k \sum_n n \cdot p_{k,n}$.

6.3.3 Other Considerations and the Final Formulation

Additional concerns must be addressed before finalizing the linear version of the formulation. Firstly, as mentioned in Section 6.2.2, the probability of the last staffing category of a sector will not exceed 1 even if the sector combines all the FPAs. The statistical models define the boundary between two adjacent staffing categories, but not the upper bound of the last category. This may result in a super sector that can accommodate all the demands. A conservative approach is to set up a realistic upper bound on either the number of FPAs that a sector can be assigned (Constraint 6-21) or on the workload that a sector can accommodate (Constraint 6-22).

Upper bound on the number of FPAs that a sector could be assigned:

$$\sum_i x_{i,k} \leq NFPA_k \quad \text{for all } k \in \{1,...,K\} \tag{6-21}$$

Upper bound on the monitoring workload that a sector could accommodate:

$$\sum_i w^m{}_i \, x_{i,k} \leq NWM_k \quad \text{for all } k \in \{1,...,K\} \tag{6-22}$$

Secondly, the objective function (6-9) defined in Section 6.2.2 would work properly if all the sectors are used. When a sector is not assigned any FPA, the predicted

number of the controller positions will still be 1, i.e. $p_{k,n=1}=1$. To accurately compute the objective (6-9), the modified objective function (6-23) has a tail term z_k-1 added to compensate this overestimation. If a sector k has no assigned FPA, which means $z_k=0$ but $p_{k,n=1}=1$, then the positions used by sector k are correctly computed with the tail term, which is zero.

Similarly for the objective function (6-10), when a sector is not assigned any FPA, there are still positive values for all the categories of the controller positions. Those values are exactly the normal CDF values of the thresholds $\mu_{k,n}$ and can be computed before running optimization. Thus, the modified objective function (6-24) also adds a tail term to the original objective function (6-10) to correctly compute the expected number of controllers.

Minimize the total predicted controller shifts

$$\min \sum_{k}\sum_{n=1}^{N} n\cdot p_{k,n}+(z_k-1) \tag{6-23}$$

Minimize the total expected controller shifts

$$\min \sum_{k}\sum_{n=1}^{N} n\cdot \mathrm{P}_{k,n}+(z_k-1)\left[\Phi(\mu_{k,n=1})+\sum_{k}\sum_{n=2}^{N} n\cdot\left(\Phi(\mu_{k,n})-\Phi(\mu_{k,n-1})\right)\right] \tag{6-24}$$

Thirdly, to linearize the absolute value operator, i.e. $\left|x_{i,k}-x_{j,k}\right|$, in Constraint (6-8), an intermediate variable $o_{ij,k}$ is introduced, and Constraint (6-8) can be replaced by the following constraints:

$$\sum_{(i,i)\in E} w^{c}{}_{ij}o_{ij,k}=\overline{w^{c}}_{k} \quad \text{for all } k\in\{1,\ldots,K\} \tag{6-25a}$$

$$o_{ij,k} \geq x_{i,k} - x_{j,k} \quad \text{for all } (i,j) \in E, k \in \{1,...,K\} \tag{6-25b}$$

$$o_{ij,k} \geq x_{j,k} - x_{i,k} \quad \text{for all } (i,j) \in E, k \in \{1,...,K\} \tag{6-25c}$$

$$o_{ij,k} \leq x_{i,k} + x_{j,k} \quad \text{for all } (i,j) \in E, k \in \{1,...,K\} \tag{6-25d}$$

$$o_{ij,k} \leq 2 - (x_{i,k} + x_{j,k}) \quad \text{for all } (i,j) \in E, k \in \{1,...,K\} \tag{6-25e}$$

Lastly, there is a need to give sectors geographical meanings, although the statistical models of sector staffing are sector-specific. Unlike FPAs that have physical meanings defined by their adjacent relations, sectors are so far considered to be generic, which means there is no practicability problem for an FPA assigned to any one of sectors $1,...,K$. This would cause solution symmetry and greatly reduce the efficiency of the branch-and-bound search, which is a typical global search method used in MIP solvers. In fact, a sector consists of several default FPAs. A hypothetical yet realistic assumption is that an open sector should have at least one of the FPAs that belongs to it prior to combination. Denoting F_k the set of default FPAs of sector k, the constraint (6-26) limits the condition in which an open sector, i.e. $z_k = 1$ should have at least one of its default FPAs. With this constraint, sectors are no longer generic and have geographical meanings.

An open sector has at least one of its default FPAs:

$$\sum_{i \in \mathrm{F}_k} x_{i,k} \geq z_k \quad \text{for all } k \in \{1,...,K\} \tag{6-26}$$

Now all the nonlinear constraints are either linearized with additional variables and constraints or reformulated with the approximation technique, all the considerations regarding model applicability have been addressed. The final formulation is linear and finalized below. It is ready for solving through any MIP solver with a proper

constraint generation method. As mentioned earlier, the proposed formulation has an exponential number of cycle elimination constraints (6-5), and it is impractical to exhaustively enumerate all of them a priori. The branch-and-cut (B&C) algorithm proposed in Chapter 3.4 can be adopted. It incorporates a dynamic constraint generation process into a typical branch-and-bound (B&B) search of a MIP solver, such as Cplex or Xpress. The constraints generated are supposed to maintain the solution feasibility (i.e. no cycles) to the original problem. More details on the algorithm can be found in Section 3.4.

Alternate Objective Function 1: *Minimize the total predicted controller shifts*

$$\min \sum_{k} \sum_{n=1}^{N} n \cdot p_{k,n} + (z_k - 1) \tag{6-23}$$

Alternate Objective Function 2: *Minimize the total expected controller shifts*

$$\min \sum_{k} \sum_{n=1}^{N} n \cdot \mathrm{P}_{k,n} + (z_k - 1)\left[\Phi(\mu_{k,n=1}) + \sum_{k} \sum_{n=2}^{N} n \cdot \left(\Phi(\mu_{k,n}) - \Phi(\mu_{k,n-1}) \right) \right] \tag{6-24}$$

Subject to:

Assignment and Sector Connectivity Constraints

$$\sum_{k} x_{i,k} = 1 \quad \text{for all } i \in \{1,...,I\} \tag{6-1}$$

$$y_{ij,k} \le \begin{cases} x_{i,k} \\ x_{j,k} \end{cases} \quad \text{for all } (i,j) \in E, k \in \{1,...,K\} \tag{6-2}$$

$$x_{i,k} \le z_k \quad \text{for all } i \in \{1,...,I\},\ k \in \{1,...,K\} \tag{6-3}$$

$$\sum_{(i,j) \in E} y_{ij,k} = \sum_{i} x_{i,k} - z_k \quad \text{for all } k \in \{1,...,K\} \tag{6-4}$$

$$\sum_{(i,j)\in E(X)} y_{ij,k} \le |X| - 1 \quad \text{for any } X \subset \{1,...,I\}, |X| \ge 2, \text{ and for all } k \in \{1,...,K\} \tag{6-5}$$

Workload Aggregation Constraints

$$\sum_i w^m{}_i x_{i,k} = \overline{w^m}_k \quad \text{for all } k \in \{1,...,K\} \tag{6-6}$$

$$\sum_i w^f{}_i x_{i,k} = \overline{w^f}_k \quad \text{for all } k \in \{1,...,K\} \tag{6-7}$$

$$\sum_{(i,i)\in E} w^c{}_{ij} o_{ij,k} = \overline{w^c}_k \quad \text{for all } k \in \{1,...,K\} \tag{6-25a}$$

$$o_{ij,k} \ge x_{i,k} - x_{j,k} \quad \text{for all } (i,j) \in E, k \in \{1,...,K\} \tag{6-25b}$$

$$o_{ij,k} \ge x_{j,k} - x_{i,k} \quad \text{for all } (i,j) \in E, k \in \{1,...,K\} \tag{6-25c}$$

$$o_{ij,k} \le x_{i,k} + x_{j,k} \quad \text{for all } (i,j) \in E, k \in \{1,...,K\} \tag{6-25d}$$

$$o_{ij,k} \le 2 - (x_{i,k} + x_{j,k}) \quad \text{for all } (i,j) \in E, k \in \{1,...,K\} \tag{6-25e}$$

Constraints of Normal CDF Approximation

$$\mu_{k,n} - \boldsymbol{\beta}'_k \mathbf{W}_k = \sum_{l=1}^{8} \lambda_{k,n,l} \cdot \varepsilon_l \quad \text{for } k \in \{1,...,K\},\ n \in \{1,...,N-1\} \tag{6-11}$$

$$\phi_{k,n} = \sum_{l=1}^{8} \lambda_{k,n,l} \cdot \Phi(\varepsilon_l) \quad \text{for } k \in \{1,...,K\},\ n \in \{1,...,N-1\} \tag{6-12}$$

$$\sum_{l=1}^{8} \lambda_{k,n,l} = 1 \quad \text{for } k \in \{1,...,K\},\ n \in \{1,...,N-1\} \tag{6-13}$$

$$\sum_{l=1}^{7} \omega_{k,n,l} = 1 \quad \text{for } k \in \{1,...,K\},\ n \in \{1,...,N-1\} \tag{6-14}$$

$$\lambda_{k,n,l} \le \omega_{k,n,l} + \omega_{k,n,l-1} \quad \text{for } k \in \{1,...,K\},\ n \in \{1,...,N-1\},\ l = \{2,...,7\} \tag{6-15a}$$

$$\lambda_{k,n,l} \le \omega_{k,n,l} \quad \text{for } k \in \{1,...,K\},\ n \in \{1,...,N-1\},\ l = 1 \tag{6-15b}$$

$$\lambda_{k,n,l} \le \omega_{k,n,l-1} \quad \text{for } k \in \{1,...,K\},\ n \in \{1,...,N-1\},\ l = 8 \tag{6-15c}$$

$$\mathrm{P}_{k,n} = \phi_{k,n} - \phi_{k,n-1} \quad \text{for } k \in \{1,...,K\},\ n \in \{2,...,N-1\} \tag{6-16a}$$

$$\mathrm{P}_{k,n} = \phi_{k,n} \quad \text{for } k \in \{1,...,K\},\ n = 1 \tag{6-16b}$$

$$\mathrm{P}_{k,n} = 1 - \phi_{k,n-1} \quad \text{for } k \in \{1,...,K\},\ n = N \tag{6-16c}$$

Constraints of Argmax Linearization (for Objective Function 1)

$$q_k \ge \mathrm{P}_{k,n} \quad \text{for all } k \in \{1,...,K\},\ n \in \{1,...,N\} \tag{6-18}$$

$$1 - p_{k,n} \ge q_k - \mathrm{P}_{k,n} \quad \text{for all } k \in \{1,...,K\},\ n \in \{1,...,N\} \tag{6-19}$$

$$\sum_n p_{k,n} = 1 \quad \text{for all } k \in \{1,...,K\} \tag{6-20}$$

Upper Bound on FPAs that Form a Sector

$$\sum_i x_{i,k} \le NFPA_k \quad \text{for all } k \in \{1,...,K\} \tag{6-21}$$

$$\sum_i {w^m}_i x_{i,k} \le NWM_k \quad \text{for all } k \in \{1,...,K\} \tag{6-22}$$

Sectors and Their Default FPAs:

$$\sum_{i \in \mathrm{F}_k} x_{i,k} \ge z_k \quad \text{for all } k \in \{1,...,K\} \tag{6-26}$$

Decision Variables:

$x_{i,k}$, $y_{ij,k}$, z_k, $\omega_{k,n,l}$, $p_{k,n}$ binary;

$\overline{w^m}_k$, $\overline{w^f}_k$, $\overline{w^c}_k$, $\mathrm{P}_{k,n}$, $\lambda_{k,n,l}$, $\phi_{k,n}$, q_k, $o_{ij,k}$ continuous and positive.

6.4 Computational Experiments

6.4.1 Experiment Setup

The proposed model is applied to realistic traffic data. The enroute center at Minneapolis (ZMP) is the study object. In the previous chapter, sector-specific models have been estimated by using three traffic metrics measured on July 2-4, 2007 at typical busy periods, i.e. from 15:00 to 23:00 GMT.

There are 6 areas of operation at ZMP. We pick for illustration Area 5, which has a total of 24 FPAs in 6 sectors, due to its significance in statistical results. Fig 6-2 shows the 2-D geographical relations of the FPAs in Area 5. Taking over from the statistical models in Chapter 5, two staffing categories of a sector are considered, i.e. using one or more than one controllers because the estimation results of three staffing categories are not statistically significant. The third category might be related not only to traffic but also to other factors (Chapter 5 provides more discussion).

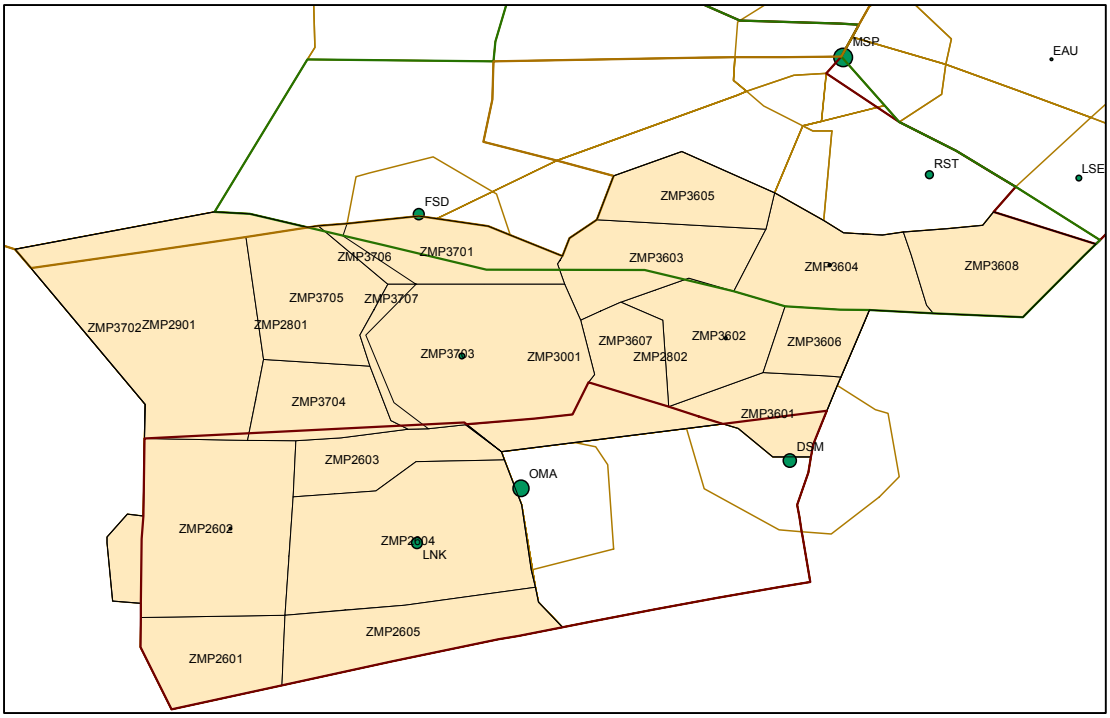


Figure 6-2 Geographical Locations of FPAs in Area 5 of ZMP

In addition, it is observed that in practice FPA combinations have certain limitations. For example, FPA3703 has only been assigned to Sectors 29, 36, and 37, and was never assigned to Sector 26, although it is geographically adjacent. These combination limitations might result from dissimilarities of traffic patterns of two sectors or from containing several altitude levels that increase difficulties for controllers. Table 6-2 summarizes such limitations observed during the selected periods in the Area 5 of ZMP and shows that some FPAs stay in their default sectors all the time while some have been assigned to up to 3 sectors.

Table 6-2 Combination Limitations during the Observation Periods in ZMP Area 5*

FPA ID**	Sector ID
ZMP2601	26
ZMP2602	26
ZMP2603	26
ZMP2604	26
ZMP2605	26
ZMP2801	28, 29
ZMP2802	28, 29, 30
ZMP2901	29
ZMP3001	29, 30
ZMP3601	36
ZMP3602	36
ZMP3603	36
ZMP3604	36
ZMP3605	36
ZMP3606	36
ZMP3607	36
ZMP3608	36
ZMP3701	37
ZMP3702	37
ZMP3703	29, 36, 37
ZMP3704	29, 36, 37
ZMP3705	29, 36, 37
ZMP3706	29, 36, 37
ZMP3707	29, 36, 37
* These restrictions are observed from 15:00 to 23:00 GMT on July 2-4, 2007. They may vary with the observation periods. ** FAP is named by center ID, sector number, and FPA number. For example, ZMP2601 represents the FPA '01' of sector '26' in 'ZMP' center.	

To examine the impact of such limitations on model performance, the proposed model can be run under the scenarios with and without additional constraints that address combination limitations. If the predictive power of the statistical models is fairly good, it is expected that with consideration of realistic limitations presented here, the optimization model would produce results close to empirical observations on controller staffing, and its performance could be validated.

Other parameters to set up are the upper bounds of Constraints (6-21) and (6-22) on how FPAs form a sector. The parameters $NFPA_k$ and NWM_k are defined as follows:

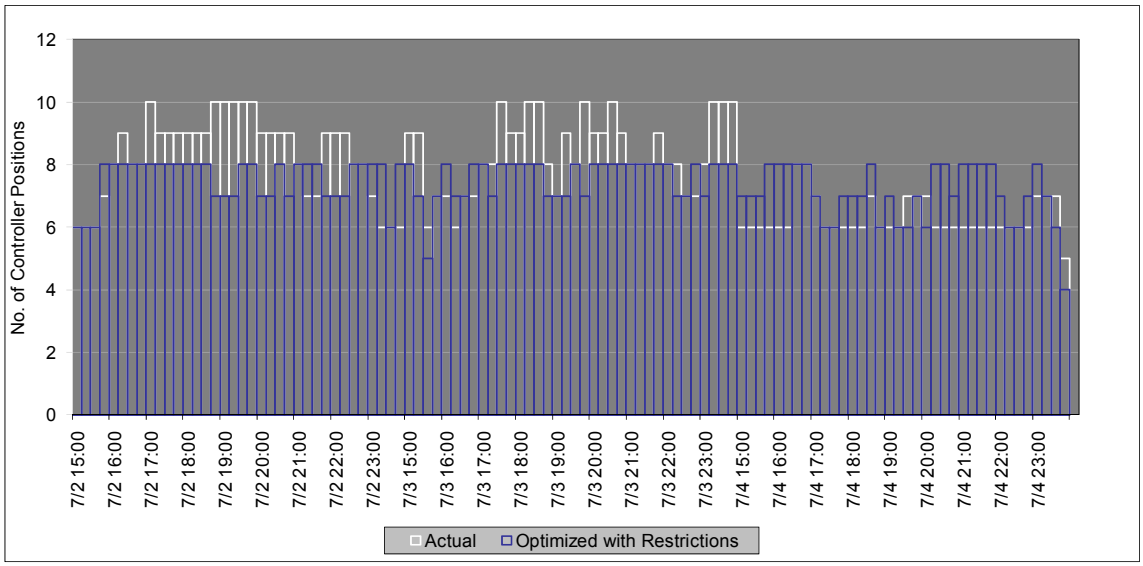
- $NFPA_k = \max\{\#\text{FPAs observed in the planning periods}\}$
- $NWM_k = \max\{\text{Monitoring workload observed in the planning periods}\}$

6.4.2 Comparison of Optimization and Real-World Results

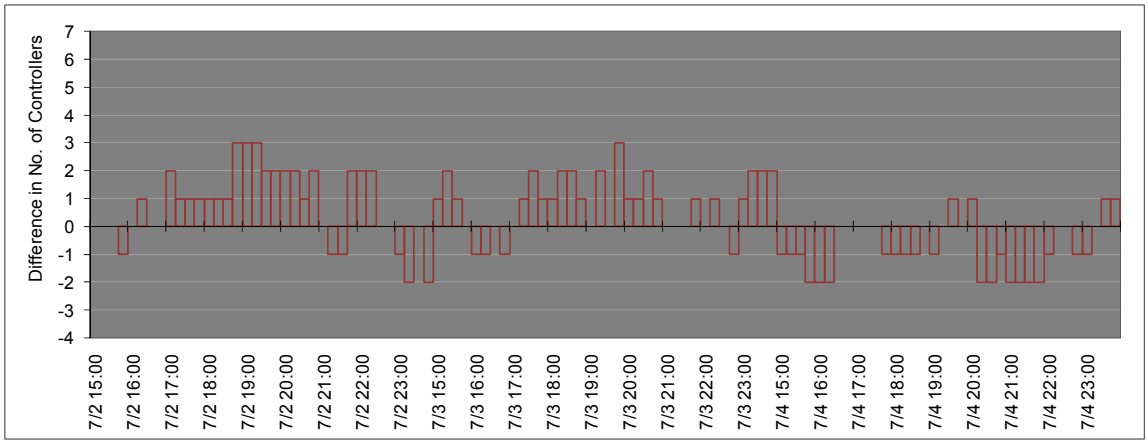
The proposed model is solved for each selected period with Xpress-Mosel ver. 2.4.1 on a Dell PowerEdge 1900 with Intel Xeon 2.66Ghz processor and 12 GB memory (single core is used, and memory is never consumed over 1 GB). The underlying network based upon FPAs in Area 5 has 24 nodes and 41 edges. Before executing the customized branch-and-cut algorithm the problem has 762 decision variables and 2051 constraints. Each instance is solved for up to 5 minutes. Over 91% of the instances can be solved to optimality within seconds.

Fig 6-3(a) compares the actual vs. predicted results with the objective of minimizing total predicted staffing and is intended to show how the predictions match with the real-world results. The white bars are the actual observations from controller staffing data. The white bars are the predicted results made jointly from the optimization model and the statistical models with historical traffic data. Fig 6-3(b) depicts the differences, where positive values mean predictions below actual outcomes. Real-world restrictions are imposed to mimic the operational environment, e.g. using pre-specified combination limitations and the upper bound for FPA counts per sector. About 30% of the selected periods are exactly matched and 40% within ±1 difference.

Arguably, the predictions made by the optimization model capture fairly the historical trend. A reason for no exact matches may be the predictive ability of the statistical models in Chapter 5, which plays an important role in estimating staffing requirements. In addition, as the optimization model minimizes the staffing levels, the positive values in Fig 6-3(b) might partially result from improving the sector combination strategy.



(a) Overlapped by Time of Day



(b) Deviation of the Predicted from Actual Values

Figure 6-3 Comparison of Actual Observation vs. Predicted Results (with the Objective Function 1 – Minimizing Total Predicted Staffing)

Beside the predicted staffing level, the expected staffing level is also of interest. Fig 6-4 compares the function values of two alternate objectives. The negative value means that the expected staffing is greater than the predicted one. Recall that the prediction is made by choosing the staffing category with the highest probability while the expected staffing is calculated by summing the product of each staffing choice and its probability. A great fraction of the comparison with negative deviation implies that the predicted staffing level is optimistic and suggests the minimum requirement of staffing to serve traffic.

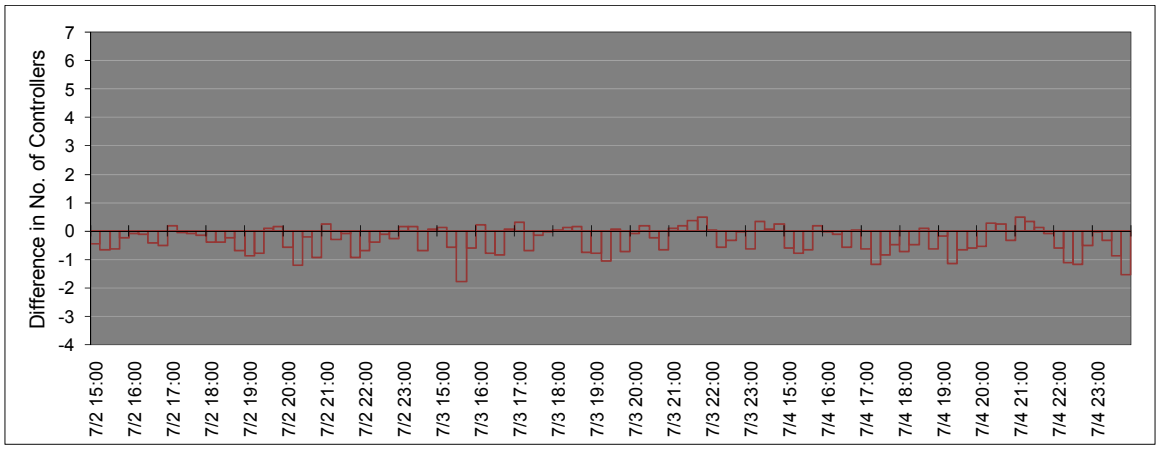
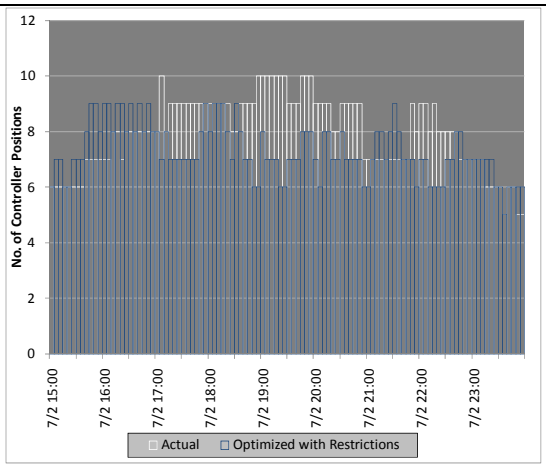


Figure 6-4 Deviation of the Value of Objective Function 2 from Objective Function 1

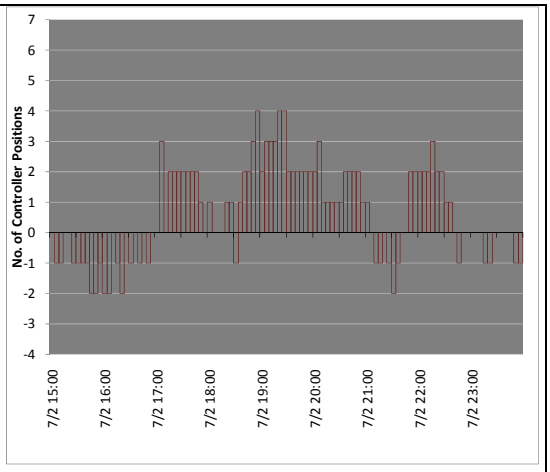
6.4.3 Results on the Period Length of 5 Minutes

In Chapter 5, the relation of traffic and staffing is also estimated on the data from 5-minute bin. The same settings are also applied to the 5-min models, and Fig 6-5 illustrates the comparison of actual observations vs. predicted results. Due to the graph resolution, the comparison results are shown for successive day. The deviations are within 4 and -2, whose range is higher than that of the 15-min models. The predictions are mostly lower than the actual observations, except for 07/04/2007.

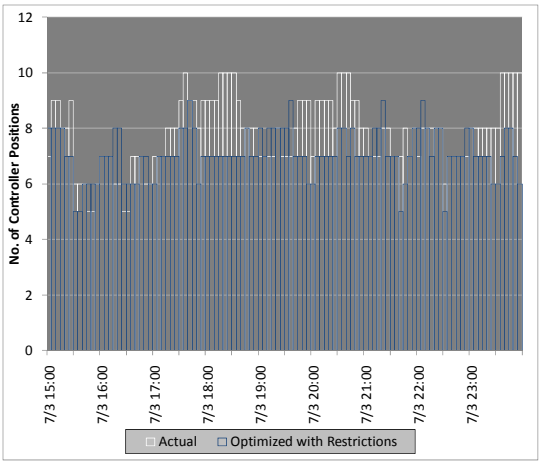
Since the experiment is set as close as possible to the real-world conditions, this underestimation trend suggests the applicability of the optimization model to the improvement of staffing efficiency.



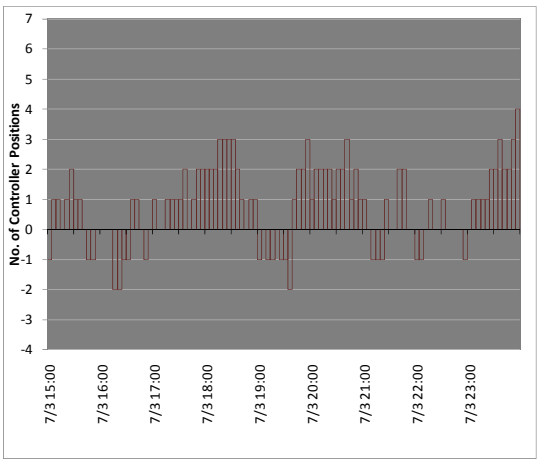
(a) Actual vs. Predicted on 07/02/2007



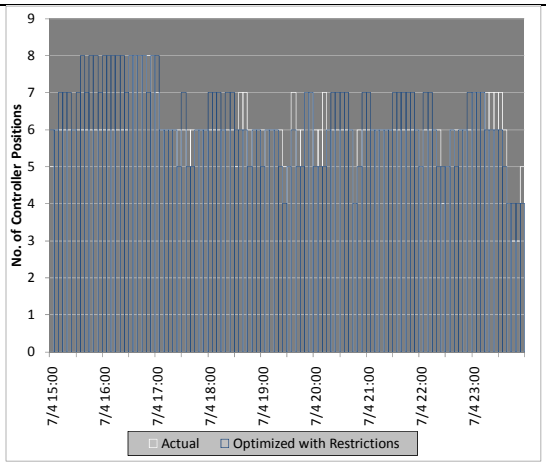
(b) Deviation on 07/02/2007



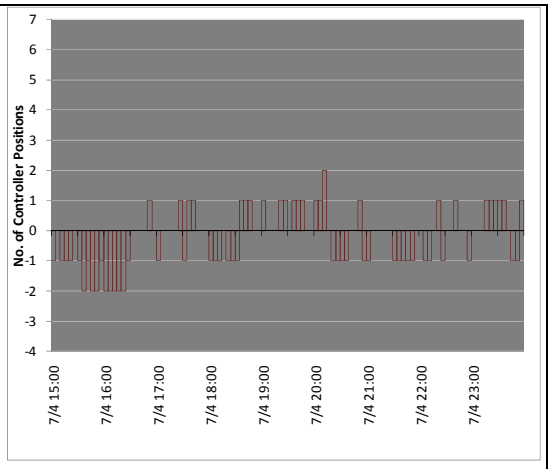
(c) Actual vs. Predicted on 07/03/2007



(d) Deviation on 07/03/2007



(e) Actual vs. Predicted on 07/04/2007



(f) Deviation on 07/04/2007

Figure 6-5 Comparison of Actual Observation vs. Predicted Results on the 5-Minute Model with the Objective Function 1

6.4.4 Relaxation Scenario Analysis

The case with the most restrictions has been considered to demonstrate the model's predictability. In Table 6-3, there are three scenarios of relaxing these restrictions to be evaluated in order to show the potential savings made with optimizing sector/FPA combinations. Each scenario is run with the objective of minimizing the predicted number of controllers, and the deviation from the optimized results before relaxation is then computed.

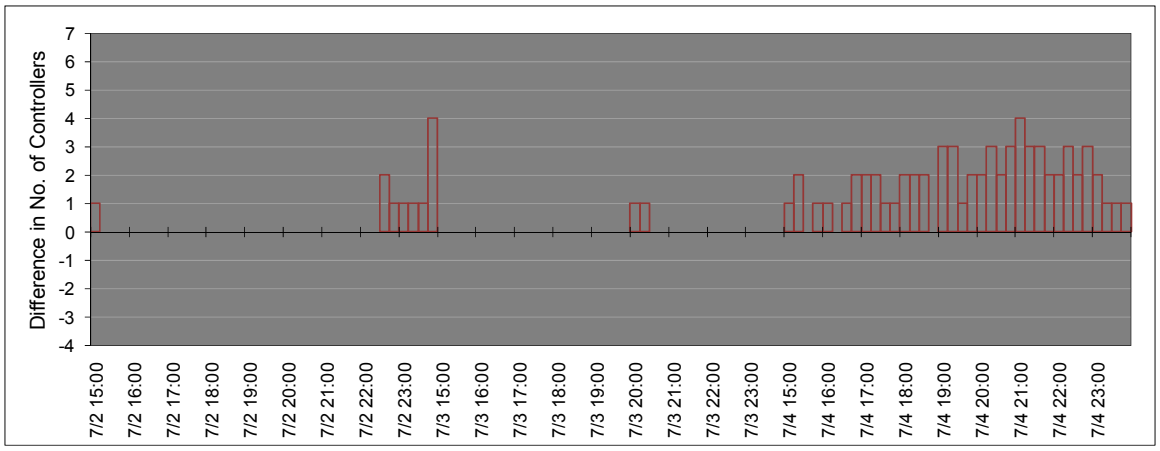
Table 6-3 Relaxation Scenarios for Sector/FPA Combination Problem

Relaxation Scenario	Constraints		
	Pre-determined Limitations	Upper Bound on FPA Counts	Upper Bound on Monitoring Workload
I	Relaxed	Imposed	Imposed
II	Imposed	Relaxed	Imposed
III	Relaxed	Relaxed	Imposed

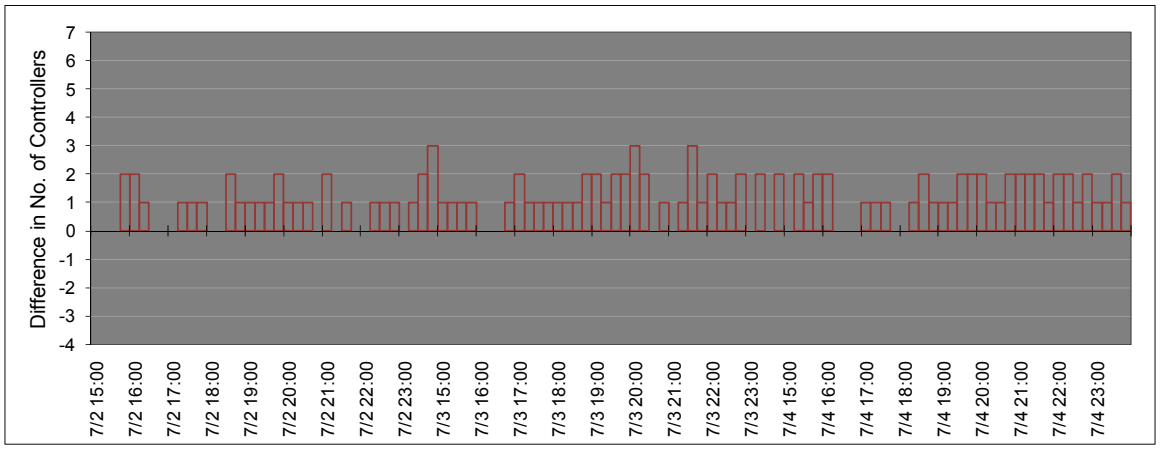
The first scenario relaxes only the pre-determined combination limitations that are observed in current practice, and its result is shown in Fig 6-6(a). The positive values mean savings and mainly fall in the selected periods on July 4, 2007. For the other two dates, the results are almost consistent with the restricted case.

The second scenario relaxes only the constraints on the number of FPAs that can be assigned to a sector. This constraint type is not unrealistic. Bloem et al. (2009) mention that the size of combined sectors should not be too large to be displayed on a scope at a resolution that allows controllers to vector aircraft. Also, the size of sectors after combinations should not be too large to stay within radio frequency coverage. With future technology improvements, the relaxation of this constraint type is foreseeable. Fig 6-6(b) shows the potential savings after relaxation.

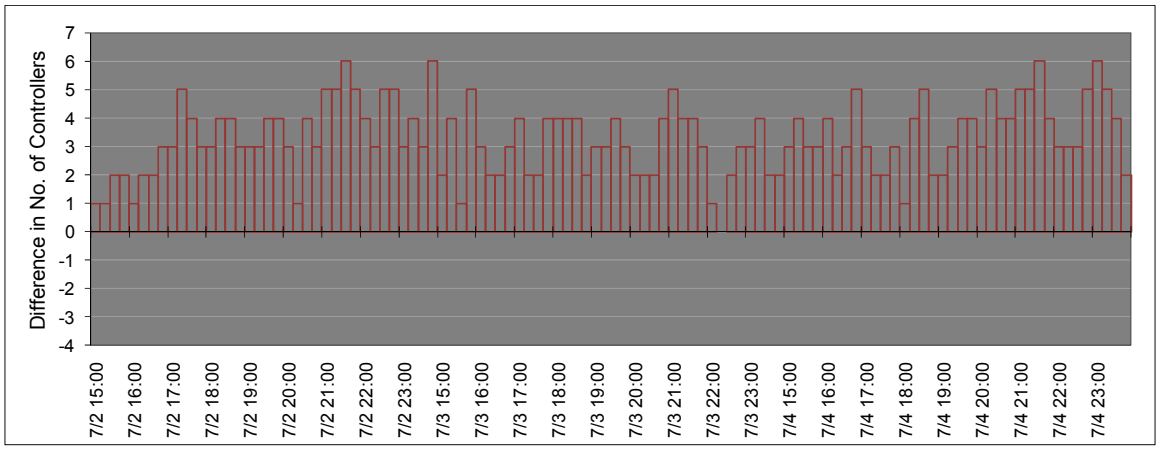
The third scenario includes both previous relaxations simultaneously while keeping the upper bound on the monitoring workload. Its savings shown in Fig 6-6(c) are the highest among all the scenarios. This is the most favorable result from sector/FPA combination optimization because no combination limitation in imposed. Since the upper bound on the monitoring workload is determined as the maximum value observed from the historical data, the scenario assumes that each sector can operate at its optimal condition, i.e. accommodate as much as traffic possible. Although safety concerns might arise when the controllers work at their full capability, the scenario itself presents a possible improvement that can be made in terms of resource efficiency.



(a) Savings of Relaxation Scenario I



(b) Savings of Relaxation Scenario II



(c) Savings of Relaxation Scenario III

Figure 6-6 Savings in Controller Resource of Relaxation Scenarios

6.5 A Time-Dependent Sector/FPA Combination Problem

The stability of sector configurations is the main concern about the sector combination problem to practitioners, especially area supervisors. Whenever there is a change of control of a sector, the new controllers must be briefed on the current situation in that sector by the current controllers. During the heat of battle, this briefing becomes costly, and frequent changes of sector combinations seem to be impractical. As most of the existing models deal with the demand of a single period, there is no guarantee of the variability in sector combinations with consecutive applications of the models to multi-period demands.

6.5.1 Modeling Time Dependency

To address the issue of time-dependent sector/FPA combinations, the proposed model is reformulated by modifying each decision variable with a time index and by adding the constraints that limit the changes of combination from periods to periods. For example, a time index t is attached to the decision variable of FPA assignment $x_{i,k}$, where $t \in \{1,...,T\}$. A time-dependent rule imposed by Constraint (6-26) is then considered: once the assignment of an FPA changes, that assignment has to last at least two consecutive periods. If a new assignment of an FPA i to some sector k happens at $t-1$, i.e. $x_{i,k,t-2} = 0$ and $x_{i,k,t-1} = 1$, then at t this FPA must stay at the sector k. On the other hand, if an FPA has been assigned to some sector for the past two consecutive periods, then the constraint is not binding in the current period.

Time-Dependency of FPA Assignment:

$$x_{i,k,t} \geq x_{i,k,t-1} - x_{i,k,t-2} \quad \text{for all } i \in \{1,...,N\}, k \in \{1,...,K\}, t \in \{3,...,T\} \qquad (6\text{-}26)$$

This relation described in Constraint (6-26), or the time-dependency (TD) constraint, connects the solutions of individual periods and governs the changes of sector/FPA combinations across the planning horizon. However, the size of the formulation inevitably grows with the number of periods under consideration, which creates a computability issue for operational purposes. In addition, the traffic forecast for the whole planning horizon is also assumed to be given and reliable; otherwise, the combination and staffing plan optimized for each period would not be applicable.

To make the TD model tractable and to consider a proper time span of a reliable demand forecast, a rolling horizon solution framework is proposed and illustrated in Fig 6-7. The idea is to decompose the planning horizon into several stages, each of which has rolling periods and remaining periods. For a stage $k \in \{1,...,K\}$, the rolling periods are actually the remaining periods in the previous stage, so the decisions for the rolling periods are inherited. The decisions for the remaining periods in the current stage are then to be optimized. As the stages progress, the decisions of each period will be determined.

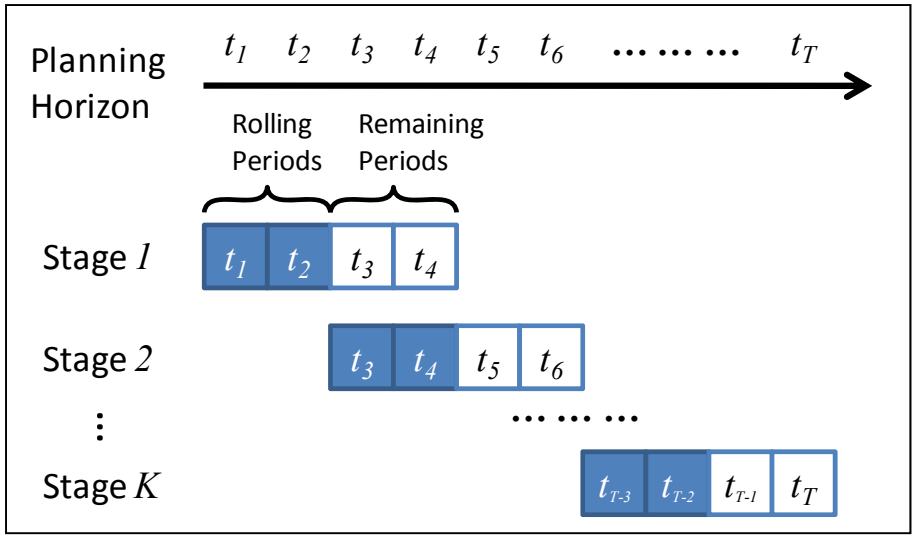


Figure 6-7 Illustration of the Rolling Horizon Framework for the Time-Dependent Problem

This rolling horizon concept has been applied to Dynamic Traffic Assignment in early transportation literature (Peeta and Mahmassani, 1995, Ran et al., 2002). For the airspace sector combination problem, the policy implication is that when the sector traffic forecast becomes reliable for the immediate periods, the decisions for the near periods can thus be based upon the ones made in earlier periods.

To integrate such a concept with the proposed model in the MIP solver, a computation framework is developed and illustrated in Fig 6-8. First, the decisions for the rolling periods in the first stage have to be initialized, so in those periods all the FPAs are assumed to stay in their default sectors. Then, the traffic data that are measured by the proposed metrics are loaded, and then the decisions about the remaining periods at the current stage are optimized in the MIP solver. If the solving process ends properly, a master solution that stores the solutions for all the periods in the planning horizon is updated with the solutions found for the remaining periods. The same procedures are repeated until the solutions of all the stages are found.

However, the TD constraint that limits the minimum stay for an FPA in a sector might result in infeasibility for some stages, especially those with a significant increase in demand. As a remedy, once the infeasibility of a stage is detected, a resolving mechanism is triggered, and the subproblem of that stage is resolved. More specifically, the resolving mechanism starts with resolving the first remaining period with TD constraints. If it still encounters the infeasibility issue, then the TD constraint is dropped, and the single-period model is solved for the first remaining period. Once the solution of the first remaining period is obtained, the mechanism is applied to the second remaining period.

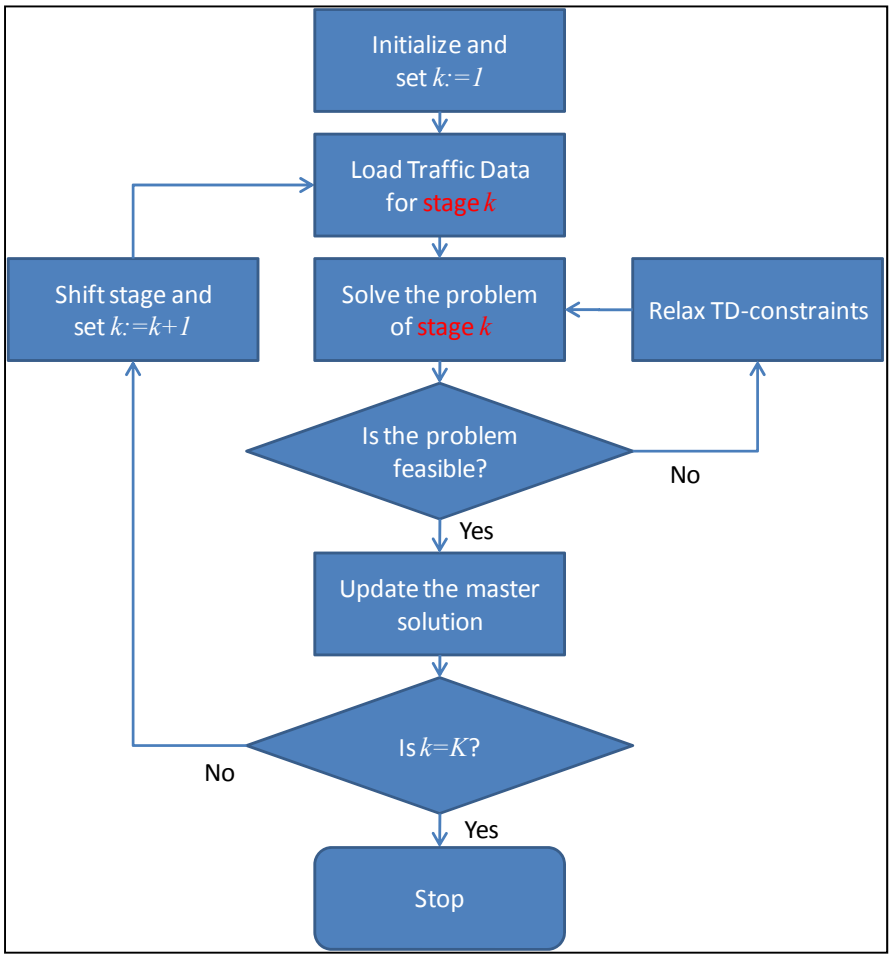


Figure 6-8 Flowchart of the Rolling Horizon Computation

6.5.2 Performance of the Time-Dependent Model

We will compare the performance of the time-dependent (TD) model with previous results. The results from two relaxation scenarios, i.e. II and III, will be used as the baseline cases. It can be anticipated that the TD model will result in using more controller resources since the minimum duration requirement limits the freedom of changing sector combinations over successive periods. In addition, to measure the assignment variability of FPAs to sectors between two consecutive periods, a metric is defined as follows:

$$\text{FPA Assignment Variability at } t = \frac{1}{2}\sum_{i,k}\left|x_{i,k,t} - x_{i,k,t-1}\right| \quad \text{, for } t = 2,...,T$$

This metric shows how many FPAs have changed sectors from $t-1$ to t by comparing the binary decisions of FPA assignment at two consecutive periods. The absolute difference is calculated and then divided by two to offset the symmetry. Its value is bounded by the number of FPAs, which is 24 in this experiment.

Fig 6-9 compares the performance of Relaxation Scenario II stated in Table 6-3 with and without the TD constraint. The negative bars illustrate that the TD model uses more controllers than the original single-period model, and thus negative savings are observed. For the variability of FPA assignment, the single-period model generally has higher values of the variability metric (the green curve) than that of the TD model (the purple curve). Noticeably, there are several spikes in the purple curve right after zero variability. Under the rolling horizon approach, as soon as the TD constraint is not in effect, a drastic change of FPA assignment occurs in order to find the sector combination that minimizes controller staffing.

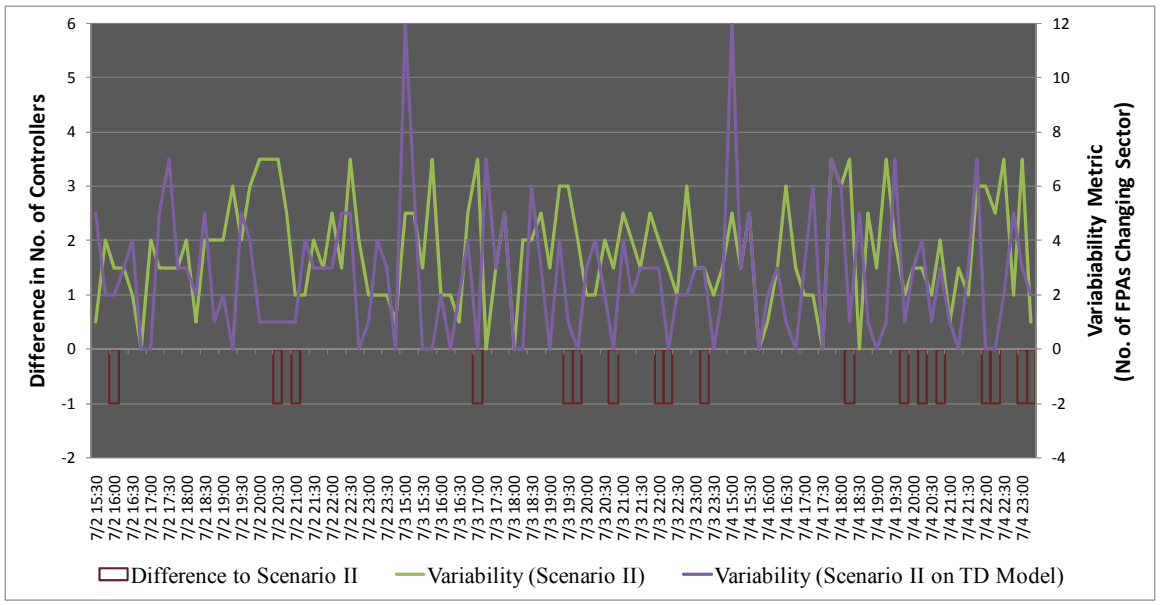


Figure 6-9 Performance Comparison of Relaxation Scenario II with and without the TD Constraint

Fig 6-10 compares the performance of Relaxation Scenario III stated in Table 6-3 with and without the TD constraint. Recall that the Scenario III relaxes the constraints on the pre-specified combination limitations and FPA counts, so it is the most optimistic scenario. The negative savings from the TD model are more significant, which means the TD model is inefficient in terms of controller staffing. Comparing the variability metric, the TD model performs significantly better than the single-period model. Since the Scenario III has the most freedom in combining FPAs/sectors, it is not surprising that the TD model reduces variability, but at the expense of using more controllers.

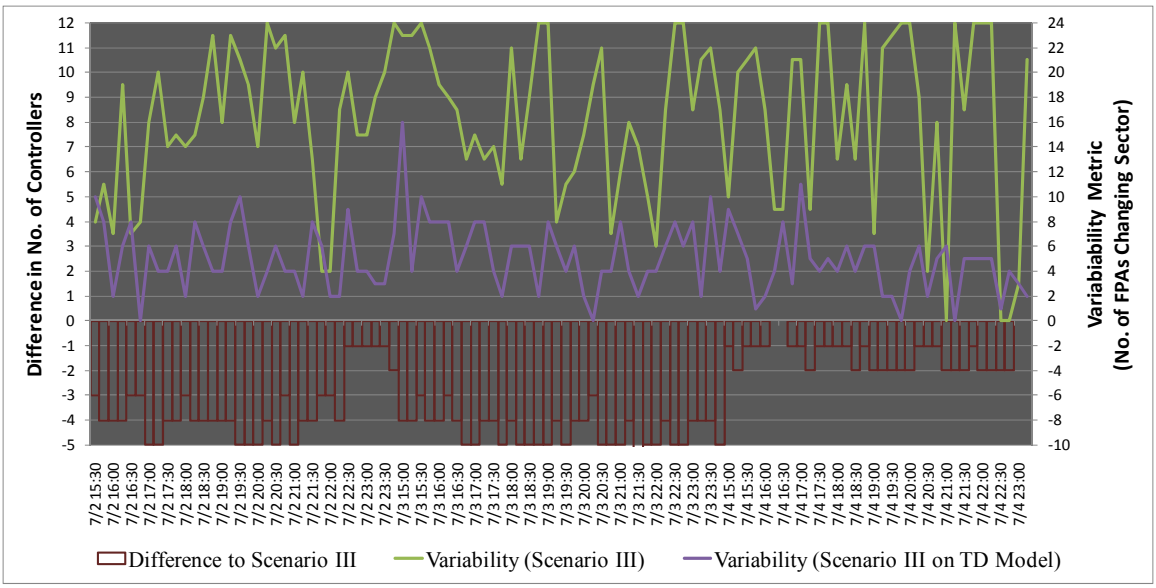


Figure 6-10 Performance Comparison of Relaxation Scenario III with and without the TD Constraint

Several descriptive statistics are summarized in Table 6-4. When the combination is restrictive, as in Scenario II, the changes in controller usage is small after adopting the TD constraint, on average increasing from 6.37 to 6.51 controllers per period, but the FPA assignment variability drops from 3.62 to 2.67 FPAs per period, which is a 26% reduction.

As Scenario III allows the least restrictive combination, the TD model has more significant impact. There is a great reduction of the FPA assignment variability, dropping from 16.36 to 5.05 FPAs per period. On the other hand, this 70% reduction is at cost of consuming more resources. Under Scenario III, the TD model results in using more than double amount of controllers, which implies using three controllers more per period to maintain the stability of sector/FPA combination.

Lastly, the success rate of the rolling horizon approach is higher for Scenario II than for Scenario III because Scenario II is more restrictive due to the pre-specified combination limitations.

Table 6-4 Statistics of Performance Comparison with or without TD Constraint

Experiment		Statistics	Relaxation Scenario	
			II	III
Without TD constraint	Controller Counts	Total	662	316
		Average	6.37	3.04
	FPA Assignment Variability	Average	3.62	16.36
		Std Dev	1.94	6.42
With TD constraint	Controller Counts	Total	677	639
		Average	6.51	6.14
	FPA Assignment Variability	Average	2.67	5.05
		Std Dev	2.41	2.73
	Successful Rate of Rolling Horizon Computation	Average	71.15%	57.69%
Difference of Controller Counts* (*negative means extra)		Total	-18	-323
		Average	-0.17	-3.11
		Percentage	2.27%	202.21%

6.7 Summary and Contributions

A mathematical model is formulated that optimizes sector/FPA combinations in order to minimize controller staffing. It is intended to help practitioners (area supervisors or air traffic managers) forecast daily or hourly staffing requirements, so that controller resources are utilized efficiently.

The model incorporates statistical estimation results that serve as the mechanism for predicting staffing levels based on traffic conditions. To solve the model with typical optimization solvers, linearization techniques are applied to approximate or transform the nonlinear objective functions and constraints. Realistic traffic and staffing data are used to evaluate the performance of the model in terms of predicting the number of

controllers needed. In the scenario analysis, the restrictions that have been imposed in the baseline model are relaxed. The potential savings in controller staffing are thus revealed by solving the problems under relaxation scenarios. As expected, the results show that the largest savings are achieved if the FPAs can be freely formed into sectors.

In modeling the time-dependent relation, the requirement of minimum duration of FPA assignment to a sector is imposed. It limits the changes of sector/FPA combinations period-by-period but at the expense of using more controllers. By maintaining stable combinations over time, we may avoid frequent changes of sector boundaries and thus reduce controller workloads, but might lose some efficiency. A possible extension is the quantitative analysis on the tradeoff between combination stability and controller staffing, so the true cost of extra burden on controller workloads due to sector boundary changes can be revealed.

Chapter 7: Conclusions and Future Extensions

7.1 Conclusions

The goals of this research are to address some neglected questions in airspace design and operation and to develop optimization models that support the planning and operational decisions required for configuring airspace. This is the first known attempt to minimize controller staffing in designing sector boundaries as well as generating sector combination schemes.

In the airspace sectorization problem, multi-period demand patterns and time-varying controller staffing are taken into consideration. The proposed model generates a set of sector boundaries that does not require frequent changes to accommodate multi-period demand patterns. Two integer program formulations, P_{STB} and P_{NFB}, are proposed to address the defined MPVC, and their equivalency is proven. Each formulation has its advantages, so the choice of formulations depends on application areas. Specifically,

- P_{STB} is based upon maintaining a spanning tree for each resulting sector whose capacity is to be optimized. P_{STB} explicitly addresses the assignment of nodes and links in the formulation, and is thus applicable to problems that require such information during optimization, such as the sector combination problem in Chapter 6. P_{STB} has constraints that grow exponentially with the problem size, so a branch-and-cut algorithm is developed for dynamically generating those constraints and solving the problem efficiently.

- P_{NFB} expands the typical network design problem by adding dummy nodes and links to represent potential sector locations and capacity values. It does not require dynamic constraint generation, so its application to a realistic size problem is still computationally tractable.

The proposed design objectives (controller cost minimization and flow alignment) are then confirmed on real traffic data from the Washington enroute center. Specifically, the performance of the proposed model is compared with a sectorization strategy that does not take demand variation into account. The numerical experiment using assumed controller capability values demonstrates that the resulting sectorization creates a design comparable to those of competing models in terms of flight dwell time and flow alignment, but saves 10%–16% controller-hours, depending on the degree of demand variation over time, which should be of great interest to air navigation service providers.

To solve a realistic-size MPVC efficiently, a large-scale neighborhood search heuristic is developed. It can find a significant improvement beyond the best solutions that are found exclusively from MIP solver's global search. The main contributions or findings from the proposed heuristic are:

- Solution quality metrics are developed that can help identify the neighborhoods for re-optimization, and their effectiveness is validated through numerical analysis.
- The impact of neighborhood sizes on heuristic performance is examined, and three neighborhood selection schemes are proposed, two of which incorporate solution quality metrics and demonstrate their effectiveness on the proposed

heuristic. Metric-based selection schemes have higher performance consistency as a result of higher improvement per success and smaller standard deviation.

- Sensitivity analysis on the quality initial solutions suggests that a good initial solution may help in finding a good improvement if the neighborhood size is large enough for escape from a local optimum.
- A parallel computation framework is designed to increase evaluations within time constraints and potentially expedite the solution evolution progress.

In the statistical analysis performed for quantifying the relation between traffic conditions and controller staffing, the ordinal regression analysis is applied on sectors of the Minneapolis enroute center and a set of statistical models for individual sectors for predicting staffing decisions is estimated. The analysis extends the previously assumed relation that controller staffing is a deterministic, step-wise function of traffic into a probabilistic, nonlinear function. It is meaningful when the length of each design period is shortened from 2 hours to 15 minutes, and the estimation results can be further incorporated into designing models for operational purposes.

Coherently with the purpose of the statistical analysis, a sector/FPA combination model is proposed for combining/splitting sectors to achieve staffing efficiency for day-to-day operations. It is intended to help practitioners (area supervisors or air traffic managers) forecast daily or hourly staffing requirements, so that controller resources are utilized efficiently while maintaining safety. The proposed formulation is based on P_{STB} and incorporates statistical estimation results that serve as the mechanism for predicting staffing levels based on traffic conditions. It is then

linearized before being solved with the branch-and-cut algorithm in an MIP solver. The numerical results show that if the FPAs can be freely formed into sectors, the largest savings in controller staffing can be achieved but at the expense of frequently changing combination schemes.

To address the combination variability over time, the proposed model is transformed into a time-dependent one. A rolling-horizon computing framework is then applied to solve the problem successively. With a defined metric of FPA assignment variability, the performance of the time-dependent model is compared with the static model. Under a restrictive combination scenario, the increase in required controllers after applying a timed-dependent model is fairly small, and the variability of FPA assignment is reduced; however, in the least restrictive combination scenario, a 70% gain in reducing variability requires doubling the number of the controllers, which implies using an average of more than three controllers per period.

7.2 Extensions

Several extensions from this dissertation seem desirable:

- Repeatedly applying a sectorization model to deal with demand changes has become a practical idea in Dynamic Airspace Configuration (DAC) research. An extended application of MPVC is that if a short period is considered, e.g. 15 or 30 minutes, the model can also be applied to support online operation. For example, a possible application setting is that given the traffic forecast for the next four 15-minute periods, a sectorization that minimizes staffing can be generated. A faster heuristic might be needed to provide real-time decision support for practitioners.

- From a practical viewpoint, there is a need for understanding the impact of frequent transition from one configuration to another on the controllers' situation awareness. Although in a controlled, simulated environment controllers could learn and work on a new configuration in a few minutes, for real-world application purposes, quantitative assessment of the workload or the risk of operational errors due to frequent transition is a must for modeling dynamically changing airspace boundaries.
- For the sector/FPA combination problem, the stability of FPA assignment concerns the practitioners. In modeling the time-dependent relation, it is found that the variability of sector/FPA combinations decrease at the expense of additional controllers. A possible extension is the quantitative analysis of the tradeoff between combination stability and controller staffing, so the true cost or extra burden on controller workload due to sector boundary changes can be revealed.
- Currently, sectorization models are studied under the assumption that a traffic forecast is given and reliable, which suggests two research directions: sectorization under demand uncertainty, and shifting traffic to achieve system efficiency.
    - To address demand uncertainty, it is possible to reformulate the proposed sectorization model into a two-stage stochastic program: the stage I variables govern the sector boundaries, and the stage II variables determine the controller staffing for various demand scenarios with the probabilities of occurrence.

- Minimizing system-wide delay through shifting traffic is a generally accepted objective of air traffic flow management (ATFM). A comprehensive model of improving ATC system efficiency may be developed by including the interaction among sectorization, controller staffing, and ATFM.

- Instead of minimizing controller staffing, the proposed airspace sectorization model or sector/FPA combination model can incorporate and allocate other ATC resources. For example, sector size or number may be bounded by radar coverage and available radio frequencies. More practical concerns can thus be added to the existing models as resource constraints or optimization objectives.

References

1. Babayev, D. A. (1997) Piece-wise Linear Approximation of Functions of Two Variables. *Journal of Heuristics*, Vol. 2, No. 4, pp313-320.
2. Ball, M. O. (2010) Heuristics Based on Mathematical Programming. To be published in *Surveys in Operations Research and Management Science*.
3. Basu, A., Mitchell, J. S. B., and Sabhnani, G. (2008) Geometric Algorithms for Optimal Airspace Design and Air Traffic Controller Workload Balancing. In Proceedings of *the 9th Workshop on Algorithm Engineering and Experiments*. San Francisco, California.
4. Bloem, M., and Kopardekar, P. (2008) Combining Airspace Sectors for the Efficient Use of Air Traffic Control Resources. In Proceedings of *AIAA Guidance, Navigation, and Control Conference and Exhibit*. Honolulu, Hawaii.
5. Bloem, M., Gupta, P., and Kopardekar, P. (2009) Algorithms for Combining Airspace Sectors. *Air Traffic Control Quarterly*, Vol. 17, No. 3, pp245-268.
6. Chatterji, G. B., Zheng, Y., and Kopardekar, P. (2008) Analysis of Current Sectors Based on Traffic and Geometry. In *AIAA Guidance, Navigation, and Control Conference and Exhibit*, Honolulu, Hawaii.
7. Conker, R., D. Moch-Mooney, W. Niedrinhaus, and B. Simmons. (2007) New Process for "Clean Sheet" Airspace Design and Evaluation. In *The 7th US/Europe Air Traffic Management R&D Seminar*, Barcelona, Spain.
8. de Berg, M., van Kreveld, M., Overmars, M. and Schwarzkopf, O. (2000) *Computational Geometry: Algorithms and Applications, Second Edition*.

Springer-Verlag, Berlin, Germany.

9. Delahaye, D., J.M. Alliot, M. Schoenauer, and J.L. Farges. (1994) Genetic Algorithms for Partitioning Air Space. In the Proceedings of *the Tenth Conference on Artificial Intelligence for Applications*, pp291–297.
10. Delahaye, D., Alliot, J., Schoenauer, M., and Farges, J. (1995) Genetic Algorithms for Automatic Regroupement of Air Traffic Control Sectors. In Proceedings of *the 4th Annual Conference on Evolutionary Programming*.
11. Delahaye, D., Schoenauer, M., and Alliot, J.-M. (1998) Airspace Sectoring by Evolutionary Computation. In Proceedings of *IEEE International Congress on Evolutionary Computation*.
12. Drew, M. (2008) Analysis of an Optimal Sector Design Method. In Proceedings of *the 27th Digital Avionics Systems Conference*. St. Paul, Minnesota.
13. Drew, M. C. (2009) A Method of Optimally Combining Sectors. In Proceedings of *the Ninth AIAA Aviation Technology, Integration and Operations Conference*.
14. Edmonds, J. (1970) *Submodular Functions, Matroids and Certain Polyhedra. Combinatorial Structures and Their Applications*. Gordon and Breach Publishing Group, New York.
15. Espinoza, D., Garcia, R., Goycoolea, M., Nemhauser, G. L., and Savelsbergh, M. W. (2008) Per-Seat, On-Demand Air Transportation Part II: Parallel Local Search. *Transportation Science*, Vol. 42, No. 3, pp279-291.
16. Federal Aviation Administration. (1997) *FAA Staffing Standards Report:*

ARTCC Radar Sector Staffing Models. Washington, D.C.

17. Federal Aviation Administration. (2006) *Submission to the United States Congress Concerning the Agency's Collective Bargaining Proposal to the National Air Traffic Controller Association*. Washington, D.C.
18. Federal Aviation Administration. (2007) *FAA's Updated Hiring Plan Matches Controller Staffing to Air Traffic*. Available at: http://www.faa.gov/news/press_releases/news_story.cfm?newsId=8251.
19. Federal Aviation Administration. (2008) *Surveillance and Broadcast Services*. Available at: http://www.faa.gov/about/office_org/headquarters_offices/ato/service_units/ enroute/surveillance_broadcast/ .
20. Federal Aviation Administration. (2008) *The Controller Work History Database (CRU-X/ART).* Retrieved on November 20, 2008.
21. Federal Aviation Administration. (2009) *Enhanced Traffic Management System (ETMS) Database*. Retrieved on June 15, 2009.
22. Federal Aviation Administration. (2010) *Order JO 7110.65T: Air Traffic Control*. Washington, D.C.
23. Federal Aviation Administration. (2010) *Order JO 7210.3W: Facility Operation and Administration*. Washington, D.C.
24. Garey, M. R., and Johnson, D. S. (1979) *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman and Company. New York, New York.
25. Gianazza, D., Alliot, J., and Granger, G. (2002) Optimal Combinations of Air Traffic Control Sectors Using Classical and Stochastic Methods. In

Proceedings of *the 2002 International Conference on Artificial Intelligence*, Las Vegas, Nevada.

26. Gianazza, D., and Alliot, J. (2002) Optimization of Air Traffic Control Sector Configurations Using Tree Search Methods and Genetic Algorithms. In Proceedings of *the 21st Digital Avionics Systems Conference*.
27. Gianazza, D. (2007) Airspace Configuration Using Air Traffic Complexity Metrics. In Proceedings of *the 7th USA/Europe Air Traffic Management R&D Seminar*, Barcelona, Spain.
28. Gianazza, D. (2008) Smoothed Traffic Complexity Metrics for Airspace Configuration Schedules. In Proceedings of *the 3rd International Conference on Research in Air Transportation*, Fairfax, Virginia.
29. Gianazza, D., Allignol, C., and Saporito, N. (2009) An Efficient Airspace Configuration Forecast. In Proceedings of *the 8th USA/Europe Air Traffic Management R&D Seminar*, Napa Valley, California.
30. Gupta, P., Bloem, M., and Kopardekar, P. (2009) An Investigation of the Operational Acceptability of Algorithm-generated Sector Combinations. In Proceedings of *AIAA Aviation Technology, Integration and Operations Conference*.
31. Holland, J. H. (1962) Outline for a Logical Theory of Adaptive Systems. *Journal of the Association for Computing Machinery*, Vol. 9, No. 3, pp297-314.
32. Hu, J., Lygeros, J., Prandini, M., and Sastry, S. (1999) Aircraft Conflict Prediction and Resolution Using Brownian Motion. In Proceedings of *the 38th*

IEEE Conference on Decision and Control.

33. Jung, J., Lee, P., Kessell, A., Homola, J., and Zelinski, S. (2010) Effect of Dynamic Sector Boundary Changes on Air Traffic Controllers. In Proceedings of *American Institute of Aeronautics and Astronautics (AIAA) Guidance, Navigation, and Control (GNC) Conference and Modeling and Simulation Technologies (MST) Conference*, Toronto, Canada.
34. Klein, A. (2005) An Efficient Method for Airspace Analysis and Partitioning based on Equalized Traffic Mass. In Proceedings of *the 6th US/Europe Air Traffic Management R&D Seminar*, Baltimore, Maryland.
35. Klein, A., Rodgers, M. D., and Kaing, H. (2008) Dynamic FPAs: A New Method for Dynamic Airspace Configuration. In Proceedings of *2008 Integrated Communications, Navigation and Surveillance Conference*, Bethesda, Maryland.
36. Kopardekar, P., Bilimoria, K., and Sridhar, B. (2007) Initial Concepts for Dynamic Airspace Configuration. In Proceedings of *the 7th AIAA Aviation Technology, Integration and Operations Conference*, Belfast, Northern Ireland.
37. Lawler, E. L. (1985) *The Traveling Salesman Problem: a Guided Tour of Combinatorial Optimization*. Wiley-Interscience Inc.
38. Lee, P., Mercer, J., Gore, B., Smith, N., Lee, K. and Hoffman, R. (2008) Examining Airspace Structural Components and Configuration Practices for Dynamic Airspace Configuration. In Proceedings of *AIAA Guidance, Navigation, and Control Conference and Exhibit*, Honolulu, Hawaii.

39. Long, J. S. (1997) *Regression Models for Categorical and Limited Dependent Variables*. Sage Publications, Inc, Thousand Oaks, California.

40. Magnanti, T. L., Mirchandani, P., and Vachani, R. (1995) Modeling and Solving the Two-Facility Modeling Network Loading Problem. *Operations Research*, Vol. 43, No. 1, pp142-157.

41. Marler, R., and Arora, J. (2004) Survey of Multi-Objective Optimization Methods for Engineering. *Structural and Multidisciplinary Optimization*, Vol. 26, No. 6, pp369-395.

42. Meyers, C., Plakosh, D., Place, P., Klein, M., and Kazman, R. (1998) *Assessment of CORBA and POSIX Designs for FAA En Route Resectorization*. Special Report CMU/SEI-98-SR-002. Carnegie Mellon University Software Engineering Institute. Pittsburgh, Pennsylvania.

43. Mitchell, J., Sabhnani, G., Krozel, J., Hoffman, R., and Yousefi, A. (2008) Dynamic Airspace Configuration Management Computational Geometry Guidance, Based on Navigation, Techniques. In Proceedings of *the Proceedings of AIAA and Control Conference and Exhibit*, Honolulu, Hawaii.

44. Mitchell, M. (1996) *An Introduction to Genetic Algorithms*. The MIT Press, Massachusetts.

45. Mogford, R., Murphy, E., and Guttman, J. (1994) Using Knowledge Exploration Tools to Study Airspace Complexity in Air Traffic Control. *The International Journal of Aviation Psychology*, Vol. 4, No. 1, pp29–45.

46. Mogford, R., Guttman, J., Morrow, S., & Kopardekar, P. (1995) *The Complexity Construct in Air Traffic Control: A Review and Synthesis of the*

Literature. DOT/FAA/CT-TN-95/22. FAA Technical Center, Atlantic City, New Jersey.

47. Nagelkerke, N. (1991) A Note on a General Definition of the Coefficient of Determination. *Biometrika*, Vol. 78, No. 3, pp691-692.
48. Peeta, S. and Mahmassani, H.S. (1995) Multiple User Classes Real-Time Traffic Assignment for On-Line Operations: A Rolling Horizon Solution Framework. *Transportation Research Part C*, Vol. 3, No. 2, pp83-98.
49. Prandini, M., Hu, J., Lygeros, J., and Sastry, S. (2000) A Probabilistic Approach to Aircraft Conflict Detection. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 1, No. 4, pp199-220.
50. Ran, B., Lee, D., and. Shin, S. (2002) Dynamic Traffic Assignment with Rolling Horizon Implementation. *Journal of Transportation Engineering*, Vol. 128, No. 4, pp314-322.
51. Rubin, J. (1973) A Technique for the Solution of Massive Set Covering Problems, with Application to Airline Crew Scheduling. *Transportation Science*, Vol. 7, pp34-48.
52. Sabhnani, G., Mitchell, J., Yousefi, A. (2010) Flow Conforming Operational Airspace Sector Design. In Proceedings of *the 10th AIAA Aviation Technology, Integration, and Operations Conference*, Fort Worth, Texas.
53. Sherali, H. D., Staats, R. W. and Trani, A. (2003) An Airspace Planning and Collaborative Decision-Making Model: Part I - Probabilistic Conflicts, Workload, and Equity Considerations. *Transportation Science*, Vol. 37, No. 4, pp434–456.

54. Sniezek, J., and Bodin, L. (2006) Using Mixed Integer Programming for Solving the Capacitated Arc Routing Problem with Vehicle/Site Dependencies with an Application to the Routing of Residential Sanitation Collection Vehicles. *Annals of Operations Research*, Vol. 144, pp33-58.

55. Sridhar, V., and Park, J. (2000) Benders-and-Cut Algorithm for Fixed-Charge Capacitated Network Design Problem. *European Journal of Operational Research*, Vol. 125, pp622–632.

56. Stein, E, Rocco, P. and Sollenberger, R. (2006) *Dynamic Resectorization in Air Traffic control: A Human Factors Perspective*. U.S. DOT/FAA Report Number DOT/FAA/TC-TN06/19. Atlantic City, New Jersey.

57. Tien, S., and Schonfeld, P. (2008) Ordinal Regression Model for Forecasting En Route Controller Requirements. In Proceedings of *the 87th Transportation Research Board Annual Meeting*, Washington, D.C.

58. Trandac, H. and Duong, V. (2003) Optimized Sectorization of Airspace with Constraints. In Proceedings of *the 5th USA/Europe Air Traffic Management R&D Seminar*, Budapest, Hungary.

59. Verlhac, C., and Manchon, S. (2001) Optimization of Opening Schemes. In Proceedings of *the 4th USA/Europe Air Traffic Management R&D Seminar*, Santa Fe, New Mexico.

60. Wolsey, L. (1998) *Integer Programming*. Wiley-Interscience Inc.

61. Xue, M. (2008) Airspace Sector Redesign Based on Voronoi Diagrams. In Proceedings of *AIAA Guidance, Navigation, and Control Conference and Exhibit*, Honolulu, Hawaii.

62. Yousefi, A. (2005) *Optimum Airspace Design with Air Traffic Controller Workload-based Partitioning*. Ph.D. Dissertation. George Mason University, Fairfax, Virginia.

63. Yousefi, A. and Donohue, G.L. (2004) Optimum Airspace Sectorization with Air Traffic Controller Workload Constraints. In Proceedings of *the 1st International Conference for Research in Air Transportation*, Zilina, Slovakia.

64. Yousefi, A., Hoffman, R., Mitchell, J., Sabhnani, G., and Krozel, J. (2007) *Enhanced Dynamic Airspace Configuration Algorithms and Concepts*. Technical Report. Metron Aviation Inc., Dulles, Virginia.